Dynamics of the Stratosphere, Mesosphere and Thermosphere

Sandford, David

Award date:
2008

Awarding institution:
University of Bath

Dynamics of the Stratosphere, Mesosphere and Thermosphere

David James Sandford

A Thesis Submitted for the Degree of Doctor of Philosophy

The University of Bath

Department of Electronic and Electrical Engineering

September 2008

“Adventure is just bad planning.”

“Eventyr er bare resultatet av dårlig planlegging.”

Roald Amundsen (1872 - 1928)

Norwegian Arctic and Antarctic Explorer

Abstract

This thesis presents observations of the dynamical features of the stratosphere, mesosphere and lower thermosphere. These are made from various observational techniques and model comparisons.

A focus of the work is the two-day wave at high latitudes in the MLT region. This has revealed significant wave amplitudes in both summer and winter. However, these waves are shown to have very different origins. Using satellite data, the summertime wave is found to be the classic quasi-two-day wave which maximises at mid-latitudes in the MLT region. The wintertime wave is found to be a mesospheric manifestation of an eastward-propagating wave originating in the stratosphere and likely generated by barotropic and baroclinic instabilities in the polar night jet.

The horizontal winds from Meteor and MF radars have been used to measure and produce climatologies of the Lunar M_2 tide at Esrange in the Arctic (68°N), Rothera and Davis in the Antarctic (68°S), Castle Eaton at mid-latitude (52°N) and Ascension Island at Equatorial latitudes (8°S). These observations present the longest period of lunar semi-diurnal tidal observations in the MLT region to date, with a 16-year dataset from the UK meteor radar. Comparisons with the Vial and Forbes (1994) lunar tidal model are also made which reveal generally good agreement.

Non-migrating lunar tides have been investigated. This uses lunar tidal results from equatorial stations, including the Ascension Island (8°S) meteor radar. Also lunar tidal results from the Rothera meteor wind radar (68°S, 68°W) and the Davis MF radar (68°S, 78°E) are considered. Both of these stations are on the edge of the Antarctic continent. It is demonstrated that there are often consistent tidal phase offsets between similar latitude stations. This suggests that non-migrating modes are likely to be present in the lunar semi-diurnal tidal structure and have significant amplitudes.

Acknowledgements

The author would like to thank in no particular order, Prof. Nicholas Mitchell for his advice and support during this project and his unending commitment to reading my thesis.

Charlotte Beldon, Dora Pancheva, Peter Younger and the rest of the CSAOS group for their support during the last 3 years.

Prof. Jeff Forbes, Prof. Robert Stening and Prof. Dennis Winch for their input regarding the error analysis of lunar M_2 tidal data and discussion into non-migrating lunar tides.

Prof. Robert Vincent and Dr. Damian Murphy for running the Davis, Antarctic MF radar and for making it available for me to use. I would also like to thank Dr. Damian Murphy for his keen input into the lunar tidal studies.

The model output of the Vial and Forbes Lunar tidal model was kindly supplied via the CEDAR database.

The Aura MLS team and Dr Michael J. Schwartz, and the work done at the Jet Propulsion Laboratory (JPL), part of the California Institute of Technology (CALTECH), which was done under contract with the National Aeronautics and Space Administration (NASA).

The UK Met Office (UKMO) and the British Atmospheric Data Centre (BADC) for supplying the Stratospheric Assimilated Data.

Also the Particle Physics and Astronomy Research Council (PPARC) for funding my studentship.

And last but by no means least; I would like to thank my parents, grandparents and sister for supporting me through my studies. Special thanks go to my mother for proofreading my work.

Contents

Page

Chapter III

Chapter IV

Chapter V

Chapter VI

Chapter VII

Chapter I

I.1 The Atmosphere

In the nineteenth and early twentieth century, data from balloons revealed that the air temperature decreased with increasing height. However, only limited measurements were made at heights above 10 km. In 1902 M. Teisserene de Bort used kite and balloon measurements to conclude that the temperature stopped decreasing at a height of between 8 and 12 km (Siskind *et al*., 2000; Holton *et al*., 2003). Previously it had been assumed that the temperature decrease must continue until the absolute zero of outer space was reached. Initially this layer of the atmosphere was named the "isothermal layer" by de Bort (Gold, 1909). Later, de Bort suggested the names *stratosphere* and *troposphere*, with the boundary between the two being named the *tropopause* by Sir Napier Shaw in 1926. Ernest Gold (1909) summarised the effects of radiative absorption on the atmosphere and noted the work of Tyndall, who conducted a series of elaborate experiments on gases known to exist in the atmosphere. Large absorption was found in water vapour and carbon dioxide but practically none in molecular nitrogen and oxygen. In contrast, ozone was found to have "remarkable absorbing power". Gold concluded that this was strong evidence for the presence of a considerable quantity of ozone in the stratosphere.

Evidence of atmospheric densities derived from meteor trails, at heights near 100 km, suggested that it was an order of magnitude denser than one would expect if the temperature profile remained constant from the tropopause (See discussion in Siskind *et al*., 2000). Auroral spectra calculated by Vegard from 1921-1936, also suggested temperatures as low as 220 K (Vegard, 1932). From these and other evidence it was suggested that there was a layer of increasing temperatures, known as the stratosphere, with another layer above in which the temperature decreases. In the 1950s many names were used by different people for the layers in the atmosphere. Chapman referred to the temperature increase (the upper part of the stratosphere) as the *mesoincline* and temperature decrease as the *mesodecline*, with a *mesopeak* in between, with these three layers making the *mesosphere*. However, after the 1950s

the temperature increase kept its name of the *stratosphere*, the maximum in between became known as the *stratopause* and the temperature decrease became the *mesosphere*, with a temperature minimum at the top first being called the *mesopause* by Sawyer in 1963 (Siskind *et al*., 2000).

In modern classifications, the atmosphere is divided into overlying layers according to the temperature gradient (e.g., Figure I.1). Each of the layers will now be considered in turn.

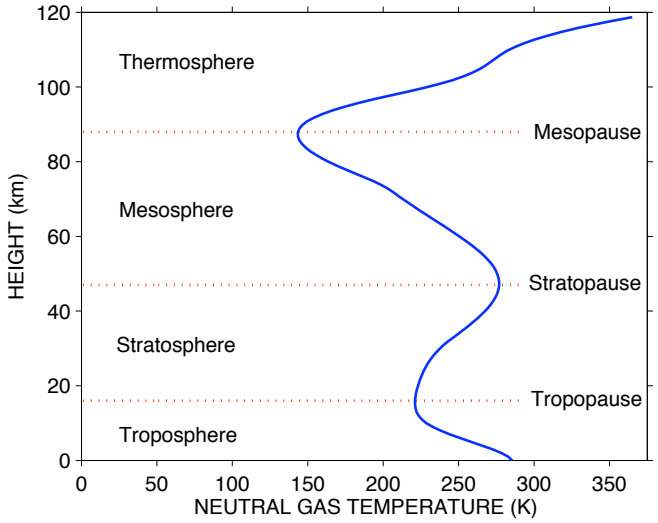


Figure I.1: Atmospheric temperature profile for July at a latitude of 55°N with an average F10.7 value of about 130 x 10^{-22}W m^{-2} Hz^{-1}, corresponding to an average to high level of solar activity (MSIS – E – 90 Model). The main regions of the atmosphere are indicated.

I.1.1 The Troposphere

The lowest layer of the atmosphere is known as the *troposphere* (literally the "turning sphere"). It is so called because this part of the atmosphere features large-scale mixing, turbulence and overturning. The troposphere contains more than 80% of the atmospheric mass, all the weather systems, most clouds and all precipitation.

In the troposphere the average temperature decreases with height at a rate of near 6.5 Kkm^{-1}. This is true up to the temperature minimum of the *tropopause*, i.e., the top of the troposphere, where a temperature inversion occurs. The height of the tropopause varies between about 9 and 16 km, being higher over the equator and lower over the poles. Even though the surface temperatures are generally higher over the equator, the difference in tropopause heights means that the temperatures over the equator

actually decrease to lower values than those over the poles. For example, a typical tropopause temperature in June is 230 K at 70°N versus 190 K at 10°N.

The top of the troposphere is a temperature minimum. This is the tropopause. The temperature of the atmosphere stops decreasing at this point due to absorption of the Sun's ultraviolet (UV) radiation. This absorption, as suggested by Gold (1909), is by the increasing concentrations of ozone (O_3) above the troposphere. Strong vertical motions near the equator raise the base of the ozone layer, forcing ozone to spread horizontally to higher latitudes. Near the poles, downward motions push the ozone to lower heights, giving the variation in tropopause height between the equator and poles. Because the tropopause is so cold it forces water to condense, meaning that very little water is able to ascend to higher levels.

I.1.2 The Stratosphere

The *stratosphere* (literally the "stratified sphere") is a region of the atmosphere which is resistant to vertical motions because it has high static stability. Therefore there is little dynamic exchange of air between the stratosphere and the troposphere below. This stability is because of its negative lapse rate (i.e., temperature increases with height). This temperature increase with increasing height in the stratosphere is due to the strong absorption of solar radiation. Figure I.2 shows the relative concentrations of some of the constituents in the atmosphere which can cause heating and cooling. The figure shows that in the stratosphere (~ 10-50 km) there is a large amount of heating due to high concentrations of ozone (O_3) and not as much cooling from carbon dioxide (CO_2) which radiates energy into space.

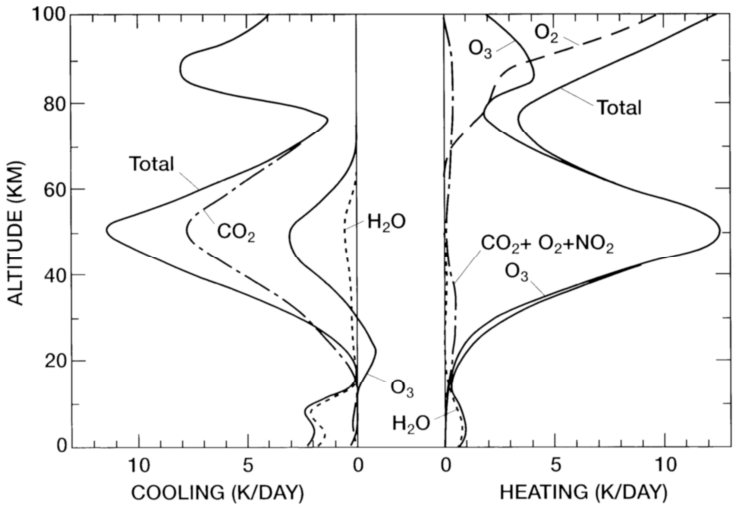


Figure I.2: The dominant heating and cooling rates of different chemical species found in the atmosphere at various heights, measured in $Kday^{-1}$ (After London, 1980).

The dynamics of the stratosphere are dominated by strong horizontal flows such as the tropical and polar-night jet streams and the quasi-biennial oscillation (QBO). The QBO is an oscillation in the zonal winds above the equator, which change direction on average every 13 months, corresponding to a period of approximately 26 months, alternating between westward and eastward winds. A dramatic phenomenon of the polar stratosphere is the *sudden stratospheric warming*, which is a rapid increase in polar stratospheric temperatures accompanied by a weakening or even reversal of the stratospheric zonal mean winds. This warming occurs because warm air from mid-latitudes is brought to higher latitudes during a distortion or break up of the polar vortex from strong planetary-wave activity. Major stratospheric warmings usually appear only in the northern hemisphere, although one was observed in the southern hemisphere in 2002.

The top of the stratosphere is the *stratopause*, at heights near 55 km where a temperature maximum occurs. This is where there is a peak in radiative heating.

I.1.3 The Mesosphere

The *mesosphere* (literally “middle sphere”) lies above the stratosphere and is a layer where the temperature decreases with increasing height. As can be seen from Figure I.2 the amount of ozone heating decreases to a minimum in the mesosphere. This is because of the depleting concentrations (measured as a fraction of the atmosphere) of atmospheric ozone in this region. Other chemical species in the mesosphere are not abundant enough to sustain the level of heating. There is however, a reduction in radiative cooling of carbon dioxide (CO_2). This is not however, enough to keep the same level of heating as in the stratosphere and mesospheric temperatures are allowed to decrease.

The mesosphere is arguably the least explored part of the atmosphere. This is because it is too high to reach with aircraft or balloons for *in situ* measurements. Similarly it is too dense for satellites orbits, so *in situ* measurements are again difficult to take. It is possible to make some *in situ* measurements with sounding rockets, falling spheres or chaff; but these only give a snapshot through the atmosphere and cannot be used from many locations. Because of these limitations, most studies of the mesosphere use either ground-based or space-based remote sensing. There are a number of different remote-sensing techniques which are now in use to measure the mesosphere. These include satellite instruments, lidar, meteor-wind radar, medium-frequency radar, incoherent-scatter radar, mesosphere-stratosphere-troposphere

radar, airglow cameras and spectrometers. All of these techniques have their individual strengths and limitations. Details of the various techniques can be found in Chapter III. The dynamics of the mesosphere and stratosphere are unique in as much as they are the only layers of the atmosphere which are strongly influenced by waves and tides. There are strong wave/mean-flow interactions, with large-scale zonal flows and tides and gravity waves of very large amplitudes.

The mesosphere hosts some strange and beautiful phenomenon. These include *noctilucent clouds*, the giant lightening discharges know as *sprites*, *meteors* and *airglow* (Figures I.3a, b and c). Noctilucent clouds are night time observations of Polar Mesospheric Clouds (Figure I.3a), the highest clouds found in the Earth's atmosphere. The height of these clouds means that they are illuminated from underneath by direct sunlight when the Sun for the observer has passed below the horizon i.e., they are brightly lit during twilight. Sprites are giant upward lightning discharges. They are launched into the mesosphere from the top of thunder storms (Figure I.3b).

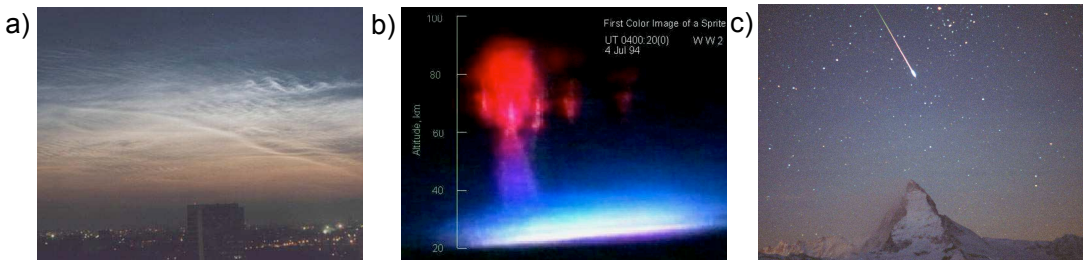


Figure I.3a: Noctilucent clouds over Moscow in 1996, b) lightning-related, stratospheric and mesospheric transient luminous events including red sprites and blue jets. c) A meteoroid falling to Earth as a bright meteor over the Alps.

The top of the mesosphere is the *mesopause* at a height of about 80 - 100 km. Figure I.4 shows the mesopause altitude as observed during the cruise of the ship MS Polarstern (Von Zahn *et al.,* 1996). It exhibits a high mesopause height (near 100 km) over the whole winter hemisphere and equatorial region. However, at 24°N the mesopause abruptly drops down towards a summer height of about 85 km.

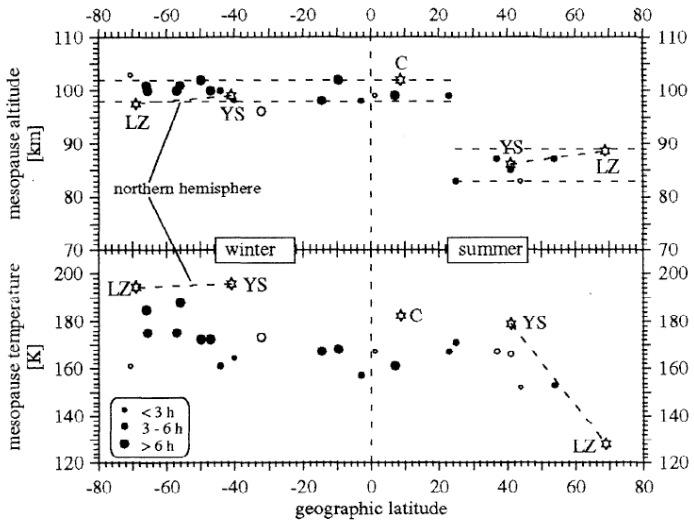


Figure I.4: Lidar observations of mesopause heights from May and June 1996 operated on board the research vessel MS Polarstern providing temperature profiles from both hemispheres (Von Zahn *et al.,* 1996).

The mesopause is the region in which extra-terrestrial material is brought into the Earth's atmosphere by ablating meteoroids which also give off the bright flashes of light sometimes visible as meteors e.g., Figures I.3c. The ablation of meteors leaves behind layers of metallic chemical species. One of these chemicals, sodium (Na), creates a layer in the atmosphere from which there is a very weak emission of light. This is caused by various processes, such as the recombination of ions which were photo-ionised by the Sun during the day or luminescence caused by cosmic rays striking the atmosphere. This is known as an airglow layer. The mesopause reaches the lowest naturally occurring temperatures found on Earth, as cold as about 125 K in polar summer.

I.1.4 The Thermosphere

The *thermosphere* (literally the "hot sphere") is the layer above the mesopause that extends out to heights of ~ 500 – 1000 km. The thermosphere comprises only a tiny fraction of the atmosphere's mass, less than 0.01% of the total. Its chemical composition changes significantly with height, as can be seen in Figure I.5. Above about 100 km molecular diffusion influences the atmospheric chemical species, separating them according to mass (i.e., in the lower thermosphere, molecular Nitrogen dominates, whereas higher up, atomic oxygen becomes the dominant constituent).

There are strong diurnal changes in the temperature of the thermosphere and the temperature also varies strongly with changing solar activity. In the thermosphere, it can reach more than 1000°C. These high temperatures result from the absorption of short-wavelength solar radiation by atomic oxygen and nitrogen. This short-wave radiation also ionises the atmosphere and leads to the formation of the *ionosphere*. The ionosphere is embedded within the neutral atmosphere.

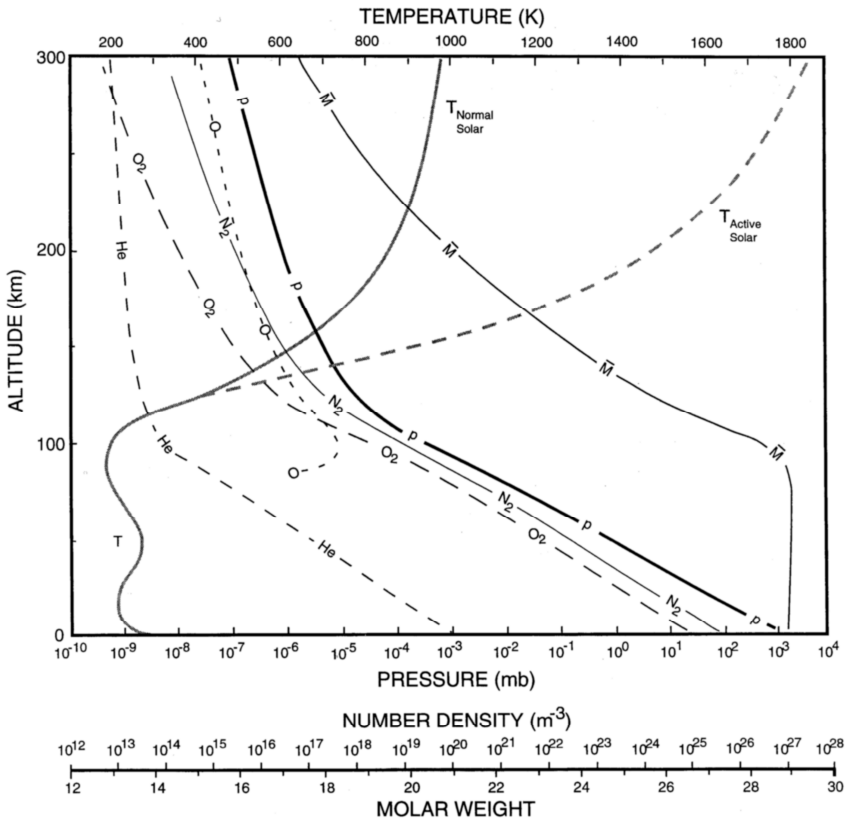


Figure I.5: Number density of atmospheric constituents (Helium (He), Atomic Oxygen (O), Molecular Oxygen (O_2) and Molecular Nitrogen (N_2)), mean molar weight ($\bar{M}$), temperature (T) and Global-mean pressure (p) of the atmosphere with height (After Salby 1996, from the U.S. Standard Atmosphere).

I.1.5 The Exosphere

The *exosphere* (literally the "outer sphere") is the outer-most layer of the atmosphere. This region is the transition between the atmosphere and space. Its lower boundary is the *thermopause / exobase* and it extends out to about 10,000 km (Dzubenko *et al.*, 2003). This is the only layer from which atmospheric gases, atoms and molecules can escape the Earth's atmosphere.

I.2 Lower, Middle and Upper Atmosphere

Another classification of layers is one that divides the atmosphere into lower, middle and upper regions. These divisions are based on the different physical processes which dominate the behaviour in each region. The *lower atmosphere* is the troposphere and is dominated by meteorological phenomena i.e., it is the region in which weather systems occur.

The *middle atmosphere* refers to the atmosphere extending from the tropopause up to about 110 km. It includes the stratosphere, mesosphere and lower thermosphere. In this region the gases are overwhelmingly neutral, so it is largely free from the influence of magnetic and electric fields. The dominant physical processes of the middle atmosphere are therefore those of fluid dynamics, including vertically propagating atmospheric waves of large amplitude.

Above the middle atmosphere is the *upper atmosphere* (above about 110 km). Here the air becomes significantly ionised and begins to be strongly influenced by the Earth's magnetic and electric fields.

I.3 The Homosphere and Heterosphere

The atmosphere can alternatively be divided into two regions, the *homosphere* and the *heterosphere*. The homosphere (literally the "same sphere") extends from the Earth's surface up to a height of about 105 km. The homosphere is a region where turbulent mixing yields a near-homogeneous composition of about 78.1% molecular nitrogen (N_2); 20.9% molecular oxygen (O_2); 0.9% argon (Ar); 0.1% carbon dioxide (CO_2) and other trace constituents. This can be seen in Figure I.5, below about 105 km, where the number density of molecular oxygen and nitrogen decreases at the same rate as the global mean pressure decreases.

The *turbopause* marks the ceiling of the homosphere and is also the height below which turbulent mixing dominates. Above the turbopause the mean free path becomes larger than the turbulent displacement of air (i.e., turbulent mixing is strongly damped).

The heterosphere (literally the "different sphere") is the Earth's atmosphere above a height of about 105 km. Within this region the molecules and atoms of the atmospheric gases each have individual scale heights.

I.4 Greenhouse Effect and Middle Atmosphere Cooling

It is clear that human activity has effects on the Earth's atmosphere. The release of chlorofluorocarbons (CFCs) has significantly increased the rate of destruction of the protective ozone layer by influencing chemical reactions in the polar stratosphere (e.g., the Antarctic ozone hole). CFC concentrations in the stratosphere have been relatively stable in recent years, and are now showing signs of decline due to the action of the Montreal Protocol (Benedick, 1991). However, a noticeable decrease in the size of the ozone hole is not expected until at least ~ 2018.

For a long while now, evidence has shown that increases in manmade greenhouse gases raise temperatures of the lower atmosphere. Greenhouse gases, such as carbon dioxide (CO_2), trap the Sun's energy, which has been absorbed by the Earth and re-emitted, which would otherwise be allowed to escape. This effectively warms the lower atmosphere and surface temperatures.

Recent modelling has also suggested that the middle atmosphere will undergo a net cooling in response to this rise in greenhouse gases. These same greenhouse gases, at a higher altitude, are able to collect energy from collisions and radiate it into space. This is due to the lower number of collisions between molecules at these heights due to the low density. This effectively cools the atmosphere. Since the late 1970's, a general cooling trend of approximately 1°C per decade has been clearly identified in the stratosphere. Increased concentrations of greenhouse gas are directly responsible for about half of the cooling and the rest comes from the loss of ozone. At present, it is predicted that at ~ 85 km, a cooling which is about 3 – 6 times greater in magnitude than the near Earth heating will occur. There is also some evidence that such a trend is already occurring faster than predicted, up to 1 Kyr^{-1}. The occurrence of Noctilucent clouds is also thought to be increasing and are being seen at lower latitudes than previously recorded. It is unclear how these changes affect the dynamics and chemistry which determine the state of the atmosphere. This area of the atmosphere has thus been dubbed a "Miner's Canary" for global change (Beig *et al.*, 2003; von Zahn, 2003; Akmaev *et al.,* 2006).

I.5 Summary

The middle atmosphere, which incorporates the Stratosphere, Mesosphere and lower Thermosphere, is an important link between the upper and lower regions of the

atmosphere. Observations in this region are made very difficult due to its conditions, i.e., too high for balloon observations and too low for satellites. This region is however, vital in identifying changes that take place, both above, below and the coupling between these regions. It is also important in ascertaining changes in the global climate.

This region of the atmosphere has several unique features which include:

- Polar Mesospheric Clouds (PMCs), the highest cloud formation found in the Earth's atmosphere, and nacreous clouds also known as Polar Stratospheric Clouds (PSCs).
- Upper-atmospheric upward lightning discharges, which are also known as "Transient Luminous Events" due to the bright flashes of light. These include sprites, blue jets, and elves.
- A direct influx of extraterrestrial material from meteoroids and particle precipitation of cosmic and solar wind sources which produces the bright streaks of Auroral emission.

This chapter introduced the atmosphere; how it is defined and how it is classified for study. Some beautiful, yet poorly understood, phenomena of the mesosphere were shown and an introduction to the middle atmosphere region, whose dynamics are dominated by winds, waves and tides, was given. The studies presented in the rest of this thesis are of the middle atmosphere, at heights of about 10 to 100 km.

Chapter II

II. Dynamics of the Middle Atmosphere

To study the dynamics of the middle atmosphere, it is important to understand a few fundamental concepts. First of all is the way in which propagating waves grow in amplitude with increasing height in the atmosphere. As an example, Figure II.1a shows temperatures and temperature perturbations measured with a Rayleigh Lidar over Aberystwyth. The results of lidar temperature measurements made over a 20 minute interval (blue line) likely consist of a background temperature profile onto which a wave is superimposed as perturbations. The second line (red) of Figure II.1a is a 4-hour average to represent a background wind state. The difference between these two lines is given in Figure II.1b, which shows temperature perturbations which are likely produced by a gravity wave. From Figure II.1b, the wave is clearly growing in amplitude with height. This shows that a small wave produced in the lower atmosphere can grow to large amplitudes by the time it reaches the mesosphere.

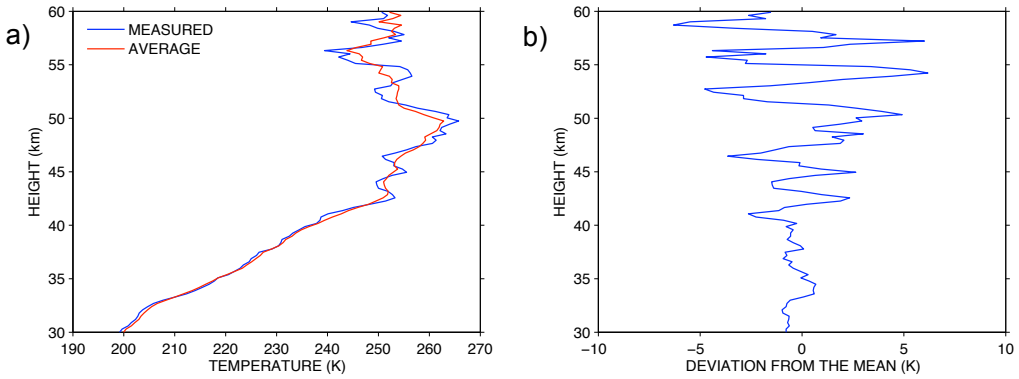


Figure II.1 a) Lidar temperature measurements from the Aberystwyth Rayleigh Lidar on the 9th November 1987. The blue line indicates measurements made over 20 minute and the red line is a 4 hour average. b) Shows the difference between the two sets of measurements in a).

The reason for this amplitude growth is that the mean density of the atmosphere decreases approximately exponentially with increasing height. If the wave is not dissipating (i.e., wave energy is conserved) this decreasing density causes wave

amplitudes to increase with height so as to conserve kinetic energy density. Kinetic energy density (KE) per unit volume is defined as $\frac{1}{2}\rho V^2$, where ρ is the air density and V is wave perturbation velocity. Therefore, as the density decreases with height, to conserve the kinetic energy, V^2 must therefore increase. This means that wave velocity amplitude will vary as $e^{z/2H}$ (where z is height and H is the scale height).

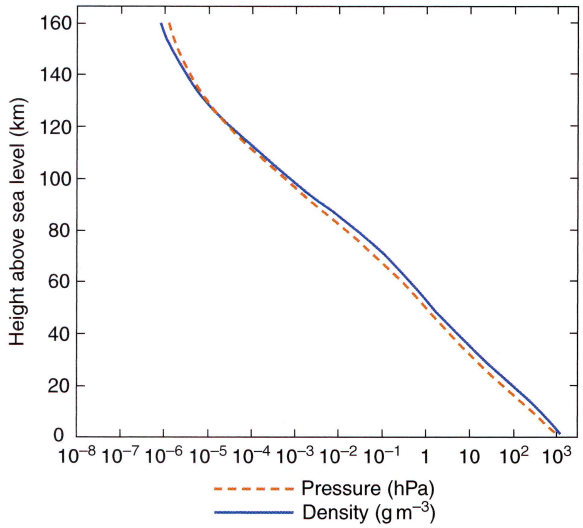


Figure II.2: The change in log density and log pressure with altitude (After Wallace and Hobbs, 2006)

The exponential decrease of density in the atmosphere is shown in Figure II.2, where pressure and density are both shown in a logarithmic scale. As the log density versus height is an approximate straight line, the density therefore falls off exponentially with height. However, wave amplitudes can become so large that they result in instabilities which act to remove energy from the wave. Such instabilities include convective and shear processes that can actually halt the exponential growth of wave amplitude. This halting of wave growth resulting from wave-generated instabilities is known as *saturation*. A wave may thus grow in amplitude until it reaches amplitudes large enough to cause instabilities and above this height, the breaking level, the wave amplitude will be constrained. This concept is illustrated in Figure II.3. The wave initially grows with height until it reaches a dissipative region in which it loses energy and therefore no longer needs to increase in amplitude.

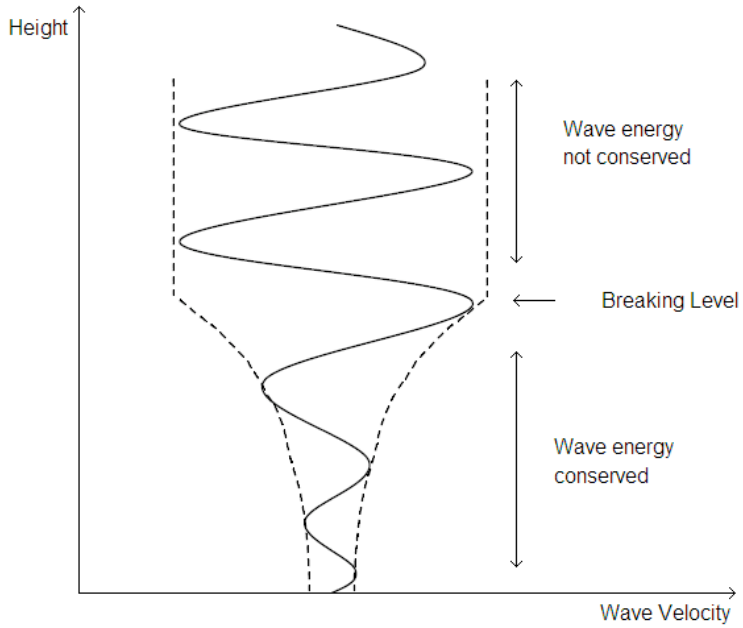


Figure II.3: Diagram showing the conservation of kinetic energy until its breaking point. In an atmosphere in which density decreases with height, to conserve the kinetic energy of a wave, its amplitude must increase until the wave is saturated.

II.1 The Primitive Equations

A set of equations which can approximate the dynamics of the middle atmosphere are known as the *primitive equations*. These primitive equations are a version of the Navier-Stokes equations that describe hydro-dynamical flow in a compressible gas which surrounds an approximately spherical rotating Earth, under the assumptions that vertical motion is much smaller than horizontal motion (hydrostasis) and that the depth of the atmosphere is small compared to the radius of the Earth. This section outlines the primitive equations and an example of their role (Wallace and Hobbs, 2006).

II.1.1 The Equation of Motion

The first of these primitive equations is the *equation of motion* and is described as the change in horizontal velocity of an air parcel of unit mass over time. It is given by:

$$\frac{du}{dt} = P + C + F \qquad \text{II.1}$$

Where u is the velocity of the air parcel, t is time, P is the pressure gradient force, C is the Coriolis force and F is a frictional force. The pressure gradient force acts to push an air parcel from areas of high pressure to areas of low pressure, where there is a pressure gradient between the two areas.

The spinning of the Earth on its axis adds the Coriolis force, which acts to push an air parcel perpendicular to its motion, acting to the right of the direction of motion in the northern hemisphere and to the left in the southern. If these are the only forces present then as the air begins to move under the pressure gradient force, it is twisted round by the Coriolis force until the air parcel is flowing parallel to the isobars and both forces are balanced. This motion is known as geostrophic flow. Figure II.4 presents a schematic diagram of the forces acting in geostrophic flow. Large parts of the stratosphere, because it is essentially frictionless, can be described well by geostrophic flow.

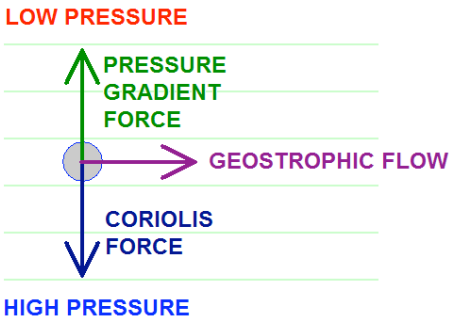


Figure II.4: Geostrophic wind created by a pressure gradient between the high pressure (e.g., at the equator) and low pressure (e.g., at the pole) for the northern hemisphere. Horizontal lines indicate isobars. As the parcel of air begins to move, due to the pressure gradient force (PGF) the Coriolis force (CF) pushes the air to the right (left) in the northern (southern) hemisphere, setting up and eastward flow with the pressure gradient and Coriolis force in balance.

The effects of friction, however, can add another force to the air parcel. An example of this in the atmosphere is friction with the surface which can become comparable to the other forces. Frictional forces have the effect of slowing the air flow as they always oppose the flow. This allows the air to converge into an area of low pressure or flow away from an area of high pressure, i.e., the air parcel is able to cross the isobars when all the forces are balanced (e.g., Andrews *et al.*, 1987). This produces a gradient wind, which is in steady state with no acceleration. An example of this is shown in Figure II.5. This type of wind is also seen in the mesosphere, where frictional drag effects come from breaking gravity waves (described further in the next section).

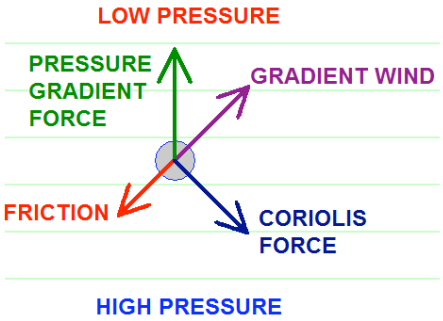


Figure II.5: Flow with a drag force (DF) included. This enables the wind to cross isobars as the pressure gradient and Coriolis force do not completely balance when the drag force is added. Horizontal lines indicate isobars.

II.1.2 The Thermodynamic Energy Equation

The second primitive equation is defined from the first law of thermodynamics. The thermodynamic energy equation is:

$$\frac{dT}{dt} = \frac{\kappa T}{p}\omega + \frac{J}{c_p} \qquad \text{II.2}$$

Where $\kappa = R/c_p$, c_P is the specific heat capacity at a constant pressure (p). $\omega = dp/dt$ which is the change in pressure with time (t). J represents the diabatic heating rate, i.e., heat sources and sinks, such as absorption of solar radiation, latent heat release and heat exchanged in atmospheric chemical reactions.

From this equation, the rate of change of an air parcel's temperature in response to the diabatic heating can be used to determine the rate of change in pressure. i.e., if an area of the atmosphere is heated by an external source, then the temperature and the pressure will rise.

II.1.2.1 Thermal Wind

The *thermal wind* is not actually a wind, but it relates the vertical shear in the geostrophic wind to a horizontal temperature gradient. It is only present in an atmosphere with horizontal gradients of temperature i.e., baroclinic. As the pressure drops off more rapidly with height for cooler air than it does for warmer air, this leads to a horizontal pressure gradient at each height level. The horizontal pressure gradient is proportional to the horizontal velocity of the wind from the equation of motion. Therefore a horizontal temperature gradient between two points will lead to a vertical

shear in the winds, with the strength of the wind becoming stronger with altitude, as long as the horizontal pressure gradient increases with height.

II.1.3 The Hypsometric Equation

The next of the primitive equations is the *hypsometric equation,* which relates the atmospheric pressure ratio to the thickness of an atmospheric layer under the assumptions of constant temperature and gravity. It is derived from the ideal gas law and the hydrostatic balance.

II.1.3.1 The Ideal Gas Law

The atmosphere can be considered as an ideal gas, which obeys the basic law,

$$pV = nRT \qquad \text{II.3}$$

Where p is the pressure; V is the volume; n is the number of moles within that volume; R is the universal gas constant and T is the temperature of the gas. This can also be extended to form a relationship between the density and its temperature and pressure. This is defined by the equation:

$$p = \frac{\rho RT}{M} \qquad \text{II.4}$$

Where ρ is the density, M is the mass of one mole of the gas.

II.1.3.2 Hydrostatic Balance

The weight of an air parcel is balanced by a difference in the pressures between the bottom and top of the air parcel, and is given by the equation:

$$\frac{dp}{dz} = -g\rho \qquad \text{II.5}$$

Where g is the acceleration due to gravity and z is the height. If this is true then the air parcel or even the entire atmosphere is said to be in hydrostatic balance.

II.1.4 The Continuity Equation

The final primitive equation, the continuity equation, is an equation which expresses conservation of mass. Matter cannot be created nor destroyed, therefore in a closed system the mass of the reactants must equal the mass of the products (Wallace and Hobbs, 2006). It is given by:

$$\frac{\delta\omega}{\delta p} = -\nabla \cdot V \qquad \text{II.6}$$

Where ω is the vertical velocity, p is the pressure and $\nabla \cdot V$ is the divergence of the horizontal wind field. Hence, a horizontal divergence ($\nabla \cdot V > 0$) of air is accompanied by a vertical compression ($\delta\omega/\delta p < 0$) and a horizontal convergence is accompanied by a vertical rarefaction.

II.2 Structure of the Stratosphere and Mesosphere in a Radiative Equilibrium Model

The heating sources and sinks of the atmosphere are generally defined by the concentrations of chemical species in the atmosphere, which absorb solar radiation and heat the air or release latent heat. This was shown in Figure I.2, with the heating and cooling rates of various atmospheric chemical species.

From the various heating rates found in Figure I.2 and assuming that the atmosphere is in a state of radiative equilibrium, it is possible to construct a global zonal mean model of the temperature expected. An atmosphere in radiative equilibrium assumes that, at all heights and latitudes, the radiative heating and cooling rates are balanced with no additional heating and cooling sources. A model of this zonally averaged temperature structure is shown in Figure II.6 for solstice conditions (Geller, 1983).

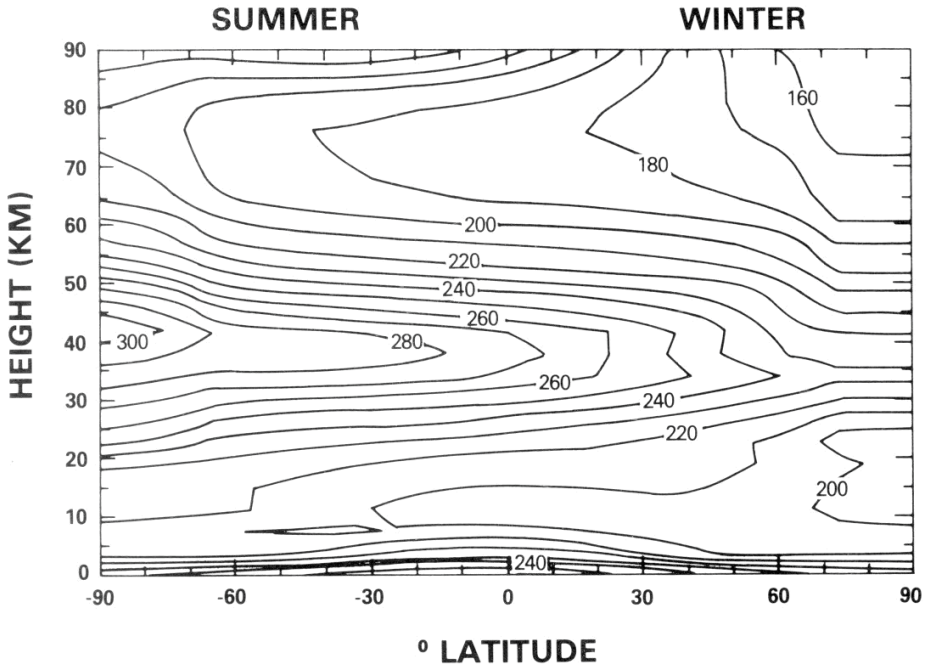


Figure II.6: The modelled global temperature structure when assuming the atmosphere is in radiative equilibrium during the solstice (Geller, 1983).

It can be seen from Figure II.6 that the summer pole has the highest temperatures at a height of ~ 43 km, as would be expected, since this pole is in 24-hour sunlight. The winter pole is predicted to have the lowest temperatures, again as might be expected, since this pole is in continual darkness so is not being heated.

In the stratosphere and mesosphere of this radiative equilibrium model, there is a gradient in the temperatures from high temperatures at the summer pole to low temperatures at the winter pole at all heights. From the thermal wind equation this temperature gradient is equivalent to a horizontal pressure gradient, with high pressures at the summer pole and low pressures at the winter pole. This pressure gradient between the north and south poles would lead to a flow evolving, where the air parcels will try to move towards the winter pole under the pressure gradient force. However, as the flow begins to move, the Coriolis force comes into play and the winds become geostrophic as the pressure gradient and Coriolis forces become equal and opposite. This leads to resultant westward geostrophic winds in the summer hemisphere and eastward geostrophic winds in the winter hemisphere. The winds flow along the isobars and remain at the same pressure level i.e., there is no vertical motion. There are therefore strong zonal winds, which are known as “jets”, and no meridional or vertical winds.

Because of the horizontal temperature gradient, a vertical shear in the winds is created, through the thermal wind equation, which leads to the geostrophic winds becoming

stronger as height increases. These Jets are then said to be “open”. Figure II.7 shows this zonal wind regime.

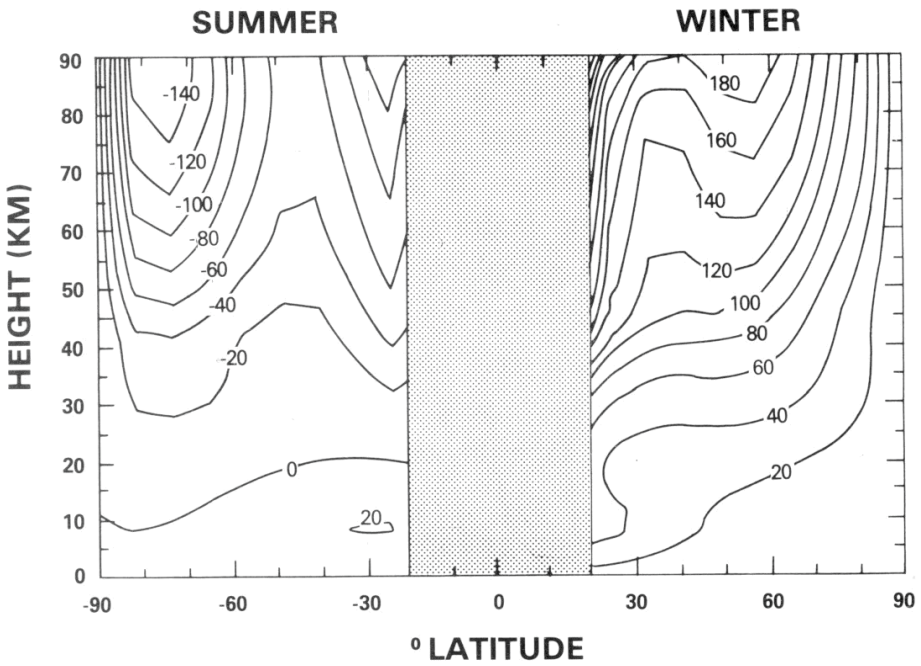


Figure II.7: Zonal mean zonal winds calculated from the radiative equilibrium temperatures model of the atmosphere. Uncapped geostrophic jets are seen at ~ 55°N travelling eastwards and ~ 75°S travelling westwards. Eastward winds are positive and westward negative. No predictions are made at the equator because of the very low Coriolis force (Geller, 1983).

II.2.1 Summary of the Radiative Equilibrium Model

We can summarise the main features of this radiative equilibrium model as follows:

- Very strong zonal winds
- Geostrophic wind increases in strength with increasing height (i.e., open jets)
- Temperature gradient with summer warmer than winter at all heights
- No meridional winds
- No vertical winds

II.3 The Observed Structure of the Stratosphere and Mesosphere

Observations have revealed a very different structure from that predicted above by the radiative equilibrium model. To illustrate these differences, we will consider the

structure of the stratosphere and mesosphere in the MSISE-90 model, which is based on observations.

The MSISE-90 model describes the neutral temperature in Earth's atmosphere from the ground to the upper thermosphere. Below a height of 72.5 km the data are primarily based on the Middle Atmosphere Program's (MAP) zonally-averaged temperature, which was also used for the CIRA-86 model (Labitzke *et al*., 1985). Below heights of 20 km the data are enhanced with averages from the National Meteorological Centre (NMC). Pitot tube, falling sphere, and grenade sounder rocket measurements from 1947 to 1972 were also used. Above 72.5 km the model takes into account data derived from mass spectrometer, space shuttle flights and incoherent scatter results (Hedin, 1991). Therefore, the MSISE-90 model, assimilates real observational data to give a reasonably reliable, zonally-averaged temperature structure.

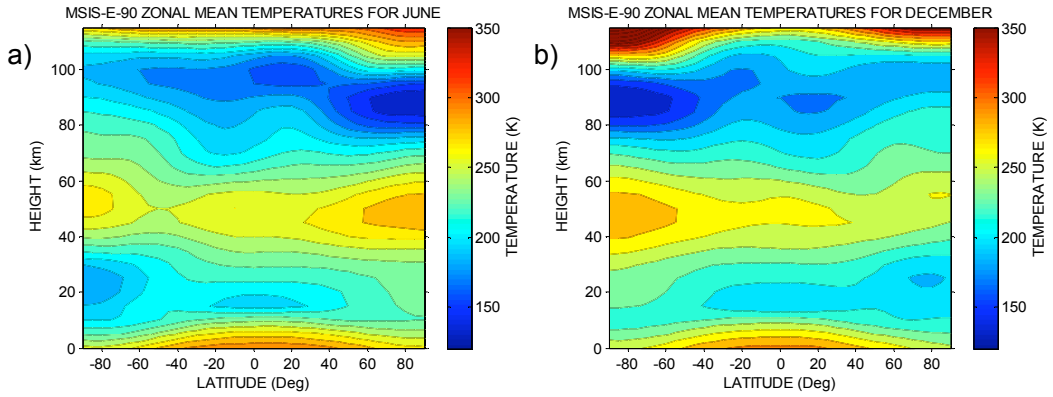


Figure II.8a, b: Zonal mean temperatures from MSISE-90 model, which uses real observational data assimilated into the model, a) for June (i.e., summer solstice) and b) for December (i.e., winter solstice).

The middle atmospheric MSIS-E-90 temperatures are presented in Figure II.8a and b for solstice conditions. Immediately noticeable is the lack of the expected north-south temperature gradient at heights above 60 km. There is however, a very cold region at mesopause heights (~ 90 km) during summer where temperatures fall to well below 140 K. Even more surprising is that the winter pole at the same height is actually *warmer* than the summer pole. For example, at a height of 90 km during June, the summer polar temperature is 133 K compared to 206 K over the winter pole. These observations differ markedly from the predictions of the radiative equilibrium model, which predicted a warmer summer mesopause.

Observations of the zonal-mean zonal wind also show significant differences from those predicted by the radiative equilibrium model. Figure II.9a and b, shows the zonal-mean zonal winds from the UARS (Upper Atmosphere Research Satellite)

Reference Atmosphere Project (URAP). The data is compiled from observations made by UARS and the Met Office stratospheric data assimilation system (Swinbank and Ortland, 2003). This data is presented because it is one of the best representations of the global-scale zonal flow of the stratosphere and mesosphere.

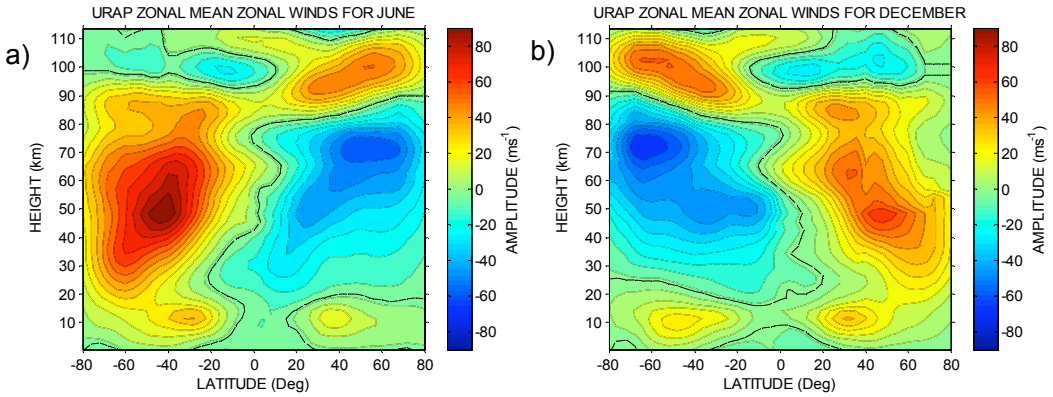
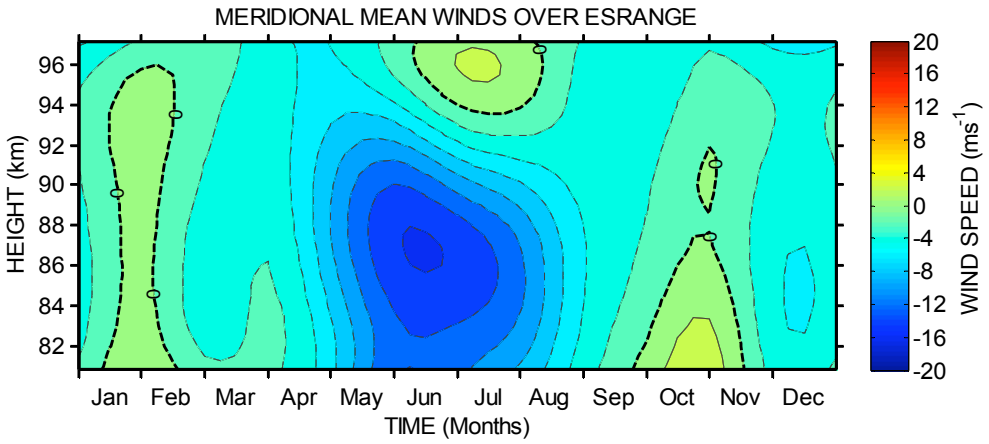


Figure II.9a, b: Zonal mean zonal winds from the URAP, which uses real observational data from the UARS satellite assimilated into the model, a) for June results (i.e., northern hemisphere summer) and b) for December results (i.e., southern hemisphere summer). Note that eastward is positive.

The zonal-mean zonal winds from Figure II.9a and b, show that there are two very strong jets which maximise at heights of between 45 and 70 km, the flow is westward in the summertime and eastward in the wintertime, as predicted. However, these jets are "closed" (i.e., their strength decreases). In fact, the jets reverse direction above the mesopause (~ 90 km). There are also two jets in the lower atmosphere which are not predicted by the radiative equilibrium model. These jets are found near to the top of the troposphere at mid latitudes and consist of eastward flows in both hemispheres.

Further observation of the mesosphere have revealed significant meridional flows Figure II.10a, b shows the mean meridional winds from the Esrange and Rothera meteor radars, located in the Arctic at 68°N and the Antarctic at 68°S, respectively. The figures present monthly-means of actual meridional wind observations over a number of years. Nearly nine years for the Esrange radar from 1999 to 2008 and over three years for the Rothera radar from 2005 to 2008. Both figures show that during the respective summer months there is a strong equatorward flow. There is also a weaker poleward flow during the winter. Therefore, a flow exists from the summer pole to the winter pole.

a)



b)

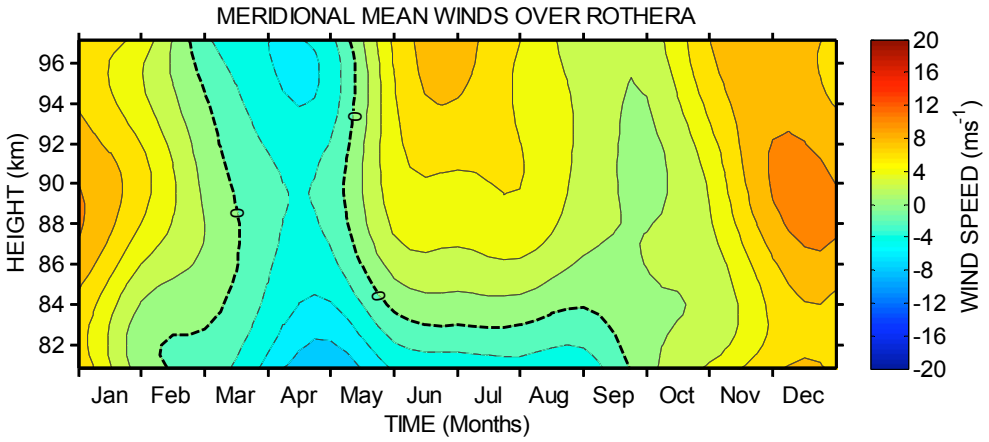


Figure II.10a, b: Composite year meridional winds from the Esrange (68°N) and Rothera (68°S) meteor wind radars averaged over 8 years from 1999 – 2008 and 3 years from 2005 – 2008, respectively. Note that northward is positive

These observations of the meridional component to the flow mean that the air converges over the winter pole and diverges away from the summer pole. From the continuity equation we know that this cannot keep happening, as the air over the winter pole needs to go somewhere and the air from the summer pole needs to be replaced. Therefore, there must be vertical motion in the polar regions. There is, therefore, an inferred up-welling of air over the summer pole and a down-welling over the winter pole. This vertical motion gives adiabatic cooling where there is an up-welling and adiabatic heating where there is a down-welling.

These heating and cooling processes, add to the radiative ones and their combination can largely account for the observed temperature structure of the stratosphere and mesosphere. The additional cooling introduced by the up-welling of the summer polar mesosphere makes this region the coldest in the terrestrial system i.e., the observed summertime mesosphere is about 90 K colder than the predicted one. In combination

the down-welling and upwelling reverse the temperature gradient between the two poles, compared to the radiative equilibrium state.

This vertical circulation is known as the *Dobson-Brewer* circulation. Figure II.11, shows the calculated movement of meridional air, from which it can be seen the transport of air occurs from the summer pole to the winter pole. The up and down welling are only of the order of a few cms^{-1} (beyond the ability of most current measurement techniques to detect).

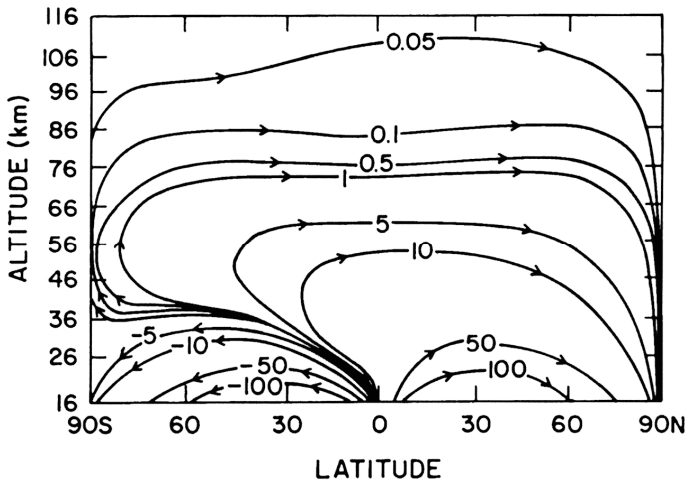


Figure II.11: Calculated mass meridional stream function in units of $kgm^{-1}s^{-1}$ during the northern hemisphere winter showing the meridional movement of the air (After Garcia and Solomon, 1983).

A consequence of this circulation is that over a summer season the entire mesosphere is deposited into the winter polar stratosphere. Curtius *et al*. (2006) found a high fraction of extra-terrestrial meteoric material within the wintertime polar vortex when taking measurements aboard Geophysica, a high-altitude Russian research aircraft. This gives experimental evidence that there is downward transport of mesospheric air into the lower stratosphere.

One important discovery made by early observations of the middle atmosphere, was finding that this temperature structure and general circulation of the mesosphere is very different to that predicted by radiative equilibrium. Therefore, there must be some other force acting on the atmosphere which has not been taken into account. From the equation of motion the only force left out, which was assumed to be small, was that of a frictional drag force. Therefore, it must be some kind of dynamical forcing. This forcing it turns out is most likely to be the effect of “wave drag”. This is believed to come from the breaking of atmospheric gravity waves, which deposits momentum directly into the atmosphere. If a drag force is added to the equation of motion calculations, this means

that the air parcel is now able to cross isobars and there is meridional flow (Andrews *et al.*, 1987).

This drag force introduced in the mesosphere in the form of gravity-wave drag from breaking waves (see section II.4.2) has the effect of closing the jets and even bringing in an eastward travelling summertime jet in the lower thermosphere. This in turn brings about the pole-to-pole temperature gradient in the mesosphere. The drag force is therefore vital in accounting for the observed circulation and temperature structure of the middle atmosphere, so the role of dynamics (i.e., Waves) in the middle atmosphere, from which this drag force evolves, is also vitally important in enhancing our understanding of the processes involved in the entire atmospheric system.

II.4 Dynamics of the Stratosphere, Mesosphere and Lower Thermosphere

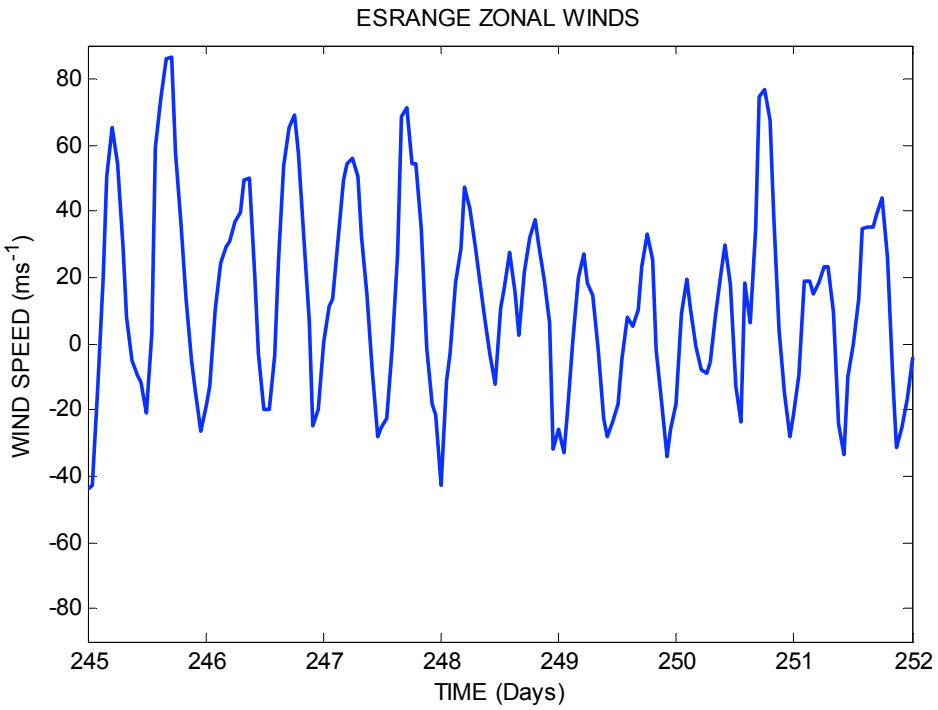
The dynamics of the stratosphere, mesosphere and lower thermosphere can be regarded as consisting of a zonally symmetric mean flow, on which are superimposed perturbations from various atmospheric waves and tides. Waves can be described as a harmonic oscillation of fluid parcels and for transverse waves the direction in which the wave propagates is perpendicular to this oscillation. These oscillations can be categorised partly in terms of the restoring force on the air parcels. This force can be produced by buoyancy in a stably stratified fluid, where the force acts to oppose a vertical displacement (e.g., gravity waves) or by the Coriolis force where the force acts to oppose a horizontal displacement (e.g., inertial waves and Rossby waves).

Global and mesoscale waves or a combination of the two (e.g., Rossby-gravity waves, Kelvin waves or tides) can be characterised by their source/propagation characteristics. Resonant modes of the atmosphere produce *free modes*. The observed quasi-2-day, 5 and 16-day large-scale planetary waves belong to this category. Waves produced by the flow over the land, convection and sudden changes in temperature are all known as *forced modes*.

The fundamental properties of various types of wave can be revealed by considering the waves' dispersion relation. The dispersion relation is a functional relation between wavenumber and frequency for a plane harmonic wave and it expresses the conditions under which a wave may propagate. It is possible to derive a dispersion relation for a particular wave, beginning with the primitive equations described earlier. These derivations usually contain a number of approximations, which are made to simplify without losing the generality of the problem.

In this chapter, we introduce three characteristic types of wave found in the middle atmosphere. These are gravity waves, tides and planetary waves. These oscillations are known to have characteristic period. Typically these are: For gravity waves between the Brunt-Väisälä period and the Coriolis period (~ 5 minutes to 15 hours at mid-latitudes); Tides have periods which are submultiples of the solar or lunar day (e.g., 12 hours); For planetary waves, the period can be anything from Coriolis period upward (> 12 hours). Figure II.12a shows a plot of the raw zonal winds from the Esrange meteor radar. In this figure there is a clear semi-diurnal tidal pattern evident, with the winds reversing direction every 12 hours. Along with this oscillation are others which are not as easy to see. The three main types of oscillation found in the middle atmosphere, described above, are shown in Figure II.12b, a spectrum of 16 years of raw winds.

a)



b)

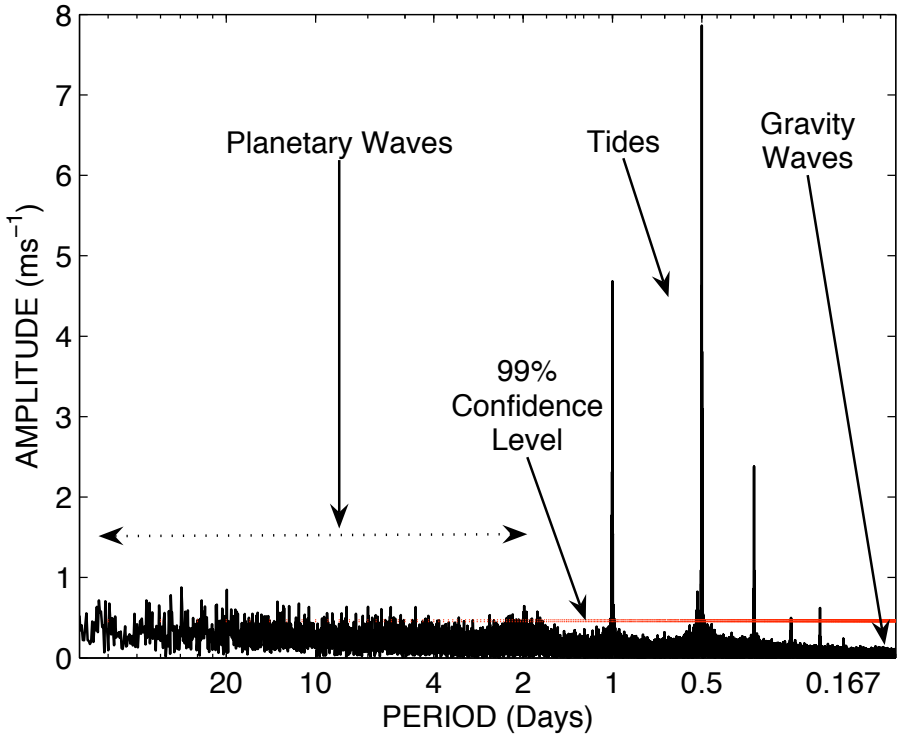


Figure II.12a, b: a) Raw zonal winds from the Esrange meteor radar. Time indicated in days from 1st January 2008. b) A Lomb-Scargle Periodogram of 16 years of zonal wind data from the UK meteor wind radar, between 1988 and 2004. Identified on the figure are planetary waves, tides and gravity waves along with a 99% significance level.

II.4.1 Brunt-Väisälä Frequency

The Brunt-Väisälä frequency also known as the buoyancy frequency is the frequency at which a vertically displaced parcel of air will oscillate in a stable environment. If the air parcel, which is initially in equilibrium with its surroundings, is displaced vertically upward, it rises adiabatically. The air parcel then has a different temperature to its surroundings due to the adiabatic expansion. If the adiabatic lapse rate is greater than the atmospheric lapse rate, as in the subadiabatic case in Figure II.13, then the air parcel (at H2 with a temperature T1) will have a lower temperature than its surroundings (T2 also at H2). From the ideal gas law in equation II.3, this air parcel will also have a greater density; therefore it will be accelerated back to its initial position. The parcel then "overshoots" its neutral position and is accelerated back the opposite way, by the same mechanism, thereby setting up an oscillation. The vertically displaced air parcel will then oscillate with simple harmonic motion about its initial position. The frequency of this oscillation is known as the Brunt-Väisälä Frequency.

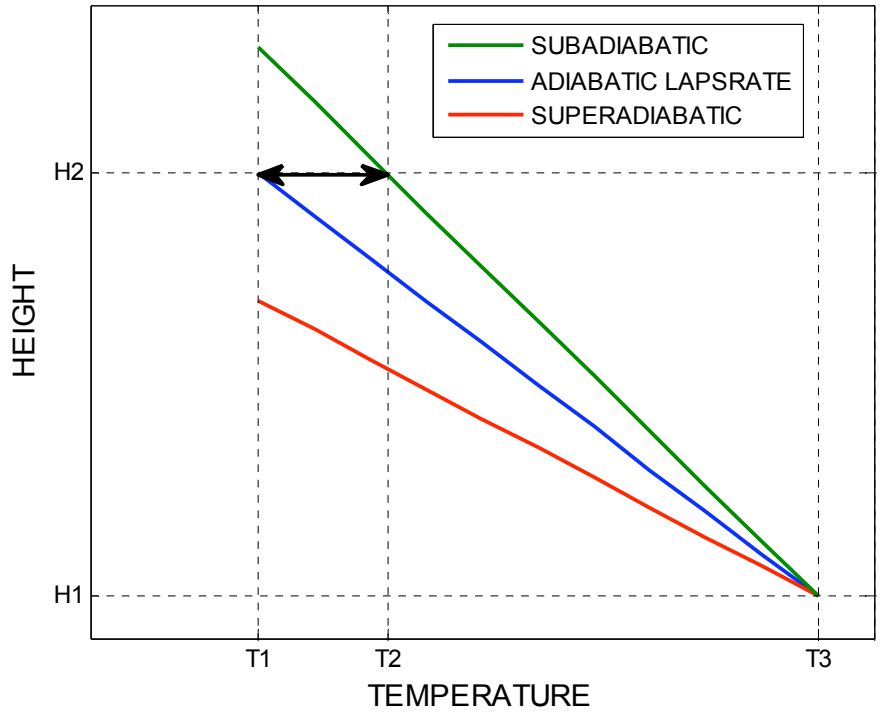


Figure II.13: The static stability of an air parcel.

If however, instead of the adiabatic lapse rate being greater than the atmospheric lapse rate, they are actually equal, there will be neutral stability and the parcel will just remain at its new height. If the atmospheric lapse rate is less than the adiabatic lapse rate (i.e., superadiabatic), the parcel in its new upper height position will have a lower density than that of its surroundings and will therefore be accelerated upwards and continue to rise. It is also true the other way round. If the air parcel initially descends, then it will be denser than its surroundings and continue to descend. In this situation the air is convectively unstable and is subject to convective overturning, which removes the unstable temperature gradient. The state of the air in this respect is known as the static stability.

II.4.2 Internal Atmospheric Gravity Waves (Buoyancy Waves)

Internal atmospheric gravity waves, commonly referred to as gravity waves, are mesoscale oscillations that arise in a stably stratified fluid when air parcels are vertically displaced. Air parcels oscillate with simple harmonic motion and buoyancy being their restoring force as described above in the Brunt-Väisälä case. These air parcels generate transverse waves by oscillating the air parcels around them, which can have horizontal and / or vertical motion. A greater understanding of the behaviour

of gravity waves is given in the following: Beer, (1974); Brekke, (1997); Jones, (2001); Fritts and Alexander (2003) and Nappo (2002).

Gravity waves which have long periods or large wavelengths > 300 km start to be affected by the Coriolis force; these waves are then referred to as inertia-gravity waves. The lowest frequency limit at which a gravity wave can oscillate is given by the inertial frequency, f, which is also the Coriolis parameter. Beyond this, the Coriolis force provides the restoring force and the waves no longer have vertical motion, these are then known as inertia waves. The Coriolis parameter is given by:

$$f = 2\Omega \sin \varphi \qquad \text{II.7}$$

Where Ω is the angular velocity of the Earth's rotation and φ is the latitude.

The longest period for a gravity wave therefore varies depending on the latitude. The highest frequency of a gravity wave is defined by the Brunt-Väisälä frequency; therefore gravity waves can only exist between these two frequencies. In the MLT region the Brunt-Väisälä frequency corresponds to a period of ~ 5 minutes. Any perturbation of a stably stratified, low-dissipation region, with frequency components between the high Brunt-Väisälä and the low Coriolis frequency can generate gravity waves.

These waves grow as they propagate upward and often reach mesospheric heights where they can become unstable and break. In breaking, they transfer their energy and momentum into the atmosphere, which drives the circulation responsible for the cold summer mesopause. This wave breaking is also known to affect other waves and tides in the middle atmosphere region (e.g. England *et al.*, 2006).

II.4.3 Planetary Waves (Rossby Waves)

Planetary waves are planetary-scale perturbation in the atmosphere that extends coherently around a full latitude circle, only existing in a rotating medium. There are two types of planetary wave, free modes and forced modes. Free modes are global normal modes of resonance in the atmosphere. They can be excited by small perturbations but they are not apparently maintained by specific forcing effects. Forced modes, however, are forced by sources such as flow over large-scale topography. Planetary waves are a dominant part of the spatial and temporal variability in the stratosphere and also make contributions at higher altitudes in the mesosphere and

lower thermosphere. They can be evident in all dynamical fields: temperature, wind, pressure and density. Most planetary waves are generated in the troposphere by the flow over large-scale topography such as mountains or the thermal contrast between the land and sea surfaces. They can also be generated *in-situ* by instabilities within the mesospheric jets. Because of the different distributions of continents in the northern and southern hemispheres, planetary waves generated in the troposphere tend to be larger in the Northern Hemisphere.

Rossby waves are a special class of planetary waves for which conservation of absolute vorticity in a barotropic atmosphere is their "restoring force". The oscillations are purely horizontal and therefore Rossby waves are a subset of inertial waves. Rossby waves are planetary in scale and have low frequencies. They arise because the Coriolis parameter varies with latitude, the so-called β effect. This links them directly with the rotation of the Earth (Charney and Drazin, 1961; Beer, 1974; Holton 1979; Salby, 1981; Andrews *et al.*, 1987).

The primitive equations can be used to derive a dispersion relation for the Rossby waves. The simplest Rossby wave solution is obtained for an atmosphere of constant density under the assumption of uniform zonal flow and no vertical motion. We therefore just consider the two dimensional $x-y$ (zonal - meridional) plane. The equations of motion in the x and y directions, respectively and a version of the continuity equation are:

$$\frac{Du}{Dt}+\frac{1}{\rho}\frac{\partial p}{\partial x}-fv=0 \qquad \text{II.8}$$

$$\frac{Dv}{Dt}+\frac{1}{\rho}\frac{\partial p}{\partial y}-fu=0 \qquad \text{II.9}$$

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \qquad \text{II.10}$$

Partially differentiate equation II.8 with $\partial/\partial y$ and II.9 with $\partial/\partial x$, then subtract. The Coriolis parameter only varies with latitude (i.e., $\partial f/\partial x=0$). It can then be written that:

$$\frac{D}{Dt}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)+f\left(\frac{\partial v}{\partial y}+\frac{\partial u}{\partial x}\right)+v\frac{\partial f}{\partial y}=0 \qquad \text{II.11}$$

The second term in II.11 is equal to zero from equation II.10. The quantity from the first term $\partial v/\partial x-\partial u/\partial y$ is known as the *relative vorticity,* ζ, here written in Cartesian

coordinates. Relative vorticity can be described as the rotation of an air parcel with respect to the coordinate system. It is thus possible to then rewrite II.11 in the form:

$$\frac{D(\zeta + f)}{Dt} = \frac{D\eta}{Dt} = 0 \qquad \text{II.12}$$

Where $\zeta + f = \eta$ is the *absolute vorticity*, ζ the relative vorticity and f is the Coriolis parameter. The absolute vorticity is the vorticity due to the rotation of the fluid itself combined with that of the rotation of the Earth. Equation II.12 shows that under the conditions of non-divergent and frictionless flow the absolute vorticity is conserved.

To solve II.12 to find the dispersion relation of a Rossby wave, we assume a linear relationship between f and y, i.e., let $f = f_0 + \beta y$, where β is the *beta plane approximation* (i.e., the rate of change of the Coriolis force from effects of the Earth's curvature). We assume a uniform zonal flow in the unperturbed situations. Linearization is used, where the variables are expanded to a background state $\overline{q}(z)$ plus a perturbation $q'(x,z,t)$, (i.e., $u = \overline{u} + u'$, $v = v'$). So that II.12, becomes:

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\left(\frac{\partial v'}{\partial x} - \frac{\partial u'}{\partial y}\right) + \beta v' = 0 \qquad \text{II.13}$$

If we assume a wave solution of the form $\exp\left[(z/2H) + i\left(k_x x + k_y y - \omega t\right)\right]$, such that $\partial/\partial t \to -i\omega$, $\partial/\partial x \to ik_x$ and $\partial/\partial y \to ik_y$. Then this gives the following dispersion relation:

$$c_x = -\frac{\omega}{k} = \overline{u} - \frac{\beta}{k_x^2 + k_y^2} \qquad \text{II.14}$$

Where c_x is the zonal phase speed of the wave. The *Rossby wave dispersion relation* of equation II.14 shows that the wave can only propagate if it has a westward phase progression relative to the zonal flow. This flow is typically only a few metres per second. Note that II.14 is a two-dimensional treatment. To extend this dispersion relation to three dimensions, without the assumption of horizontal non-divergence, the dispersion relation becomes:

$$c_x = -\frac{\omega}{k_x} = \overline{u} - \frac{\beta}{\left(k_x^2 + k_y^2 + \frac{f_0^2}{\omega_B^2}\left(k_z^2 + \frac{1}{4H^2}\right)\right)} \qquad \text{II.15}$$

Where k_z is the vertical wavenumber. For the case of stationary waves (i.e., those which do not move with respect to the surface of the Earth), the phase speed of the wave must be zero ($c_x = 0$). In this case, II.15 can be rewritten in terms of k_z as:

$$k_z^2 = \frac{\omega_B^2}{f_0^2}\left(\frac{\beta}{\overline{u}} - \left(k_x^2 + k_y^2\right)\right) - \frac{1}{4H^2} \qquad \text{II.16}$$

For vertical propagation to occur k_z^2 must be positive (real k_z), which gives the propagation condition:

$$0 < \overline{u} < \frac{\beta}{\left(k_x^2 + k_y^2\right) + \frac{f_0^2}{\omega_B}\left(\frac{1}{4H^2}\right)} \qquad \text{II.17}$$

From equation II.17, it can be seen that Rossby waves can only propagate vertically for a limited range of zonal wind speeds. If the mean zonal flow is westward or if it is eastward and exceeds a critical speed (given by the right hand side of equation II.17), the Rossby waves are external and do not propagate vertically.

For vertical propagation to occur in non-stationary Rossby waves (i.e., those which move with the respect to the surface of the Earth), the phase speed of the waves must obey the condition,

$$0 < \overline{u} - c_x < U_c \qquad \text{II.18}$$

From which,

$$U_c \equiv \frac{\beta}{k_x^2 + k_y^2 + \frac{f_0^2}{\omega_B^2}\left(\frac{1}{4H^2}\right)} \qquad \text{II.19}$$

Where U_c is the *Rossby critical velocity* (Forbes, 1995). Equations II.19 is known as the Charney-Drazin criterion. It shows that for vertical propagation of a planetary wave, its phase speed must be westward with respect to the mean wind and that the mean wind must be less than the Rossby critical velocity plus the phase speed of the wave (Charney and Drazin, 1961).

As an example of a Rossby wave, consider an air parcel which is travelling eastward (e.g., in an eastward travelling jet) in the northern hemisphere. If the air parcel is initially displaced southward, the planetary component of vorticity, f, which is represented by the Coriolis parameter in equation II.12, will decrease. Since the absolute vorticity, η, remains constant, the local relative vorticity of the air parcel, ζ,

must increase (Position 1 in Figure II.14). This results in an anti-clockwise rotation of the air mass in the air parcel, which begins to move northward. The air parcel then overshoots its equilibrium position and the planetary vorticity then begins to increase (Position 2 in Figure II.14), resulting in a decrease in the relative vorticity and the air parcel spins clockwise. It then turns south again (Position 3 in Figure II.14). The wave itself will propagate in a westward direction. Note that to satisfy the criterion II.15 and II.18 the wind must be travelling eastward for propagation to occur.

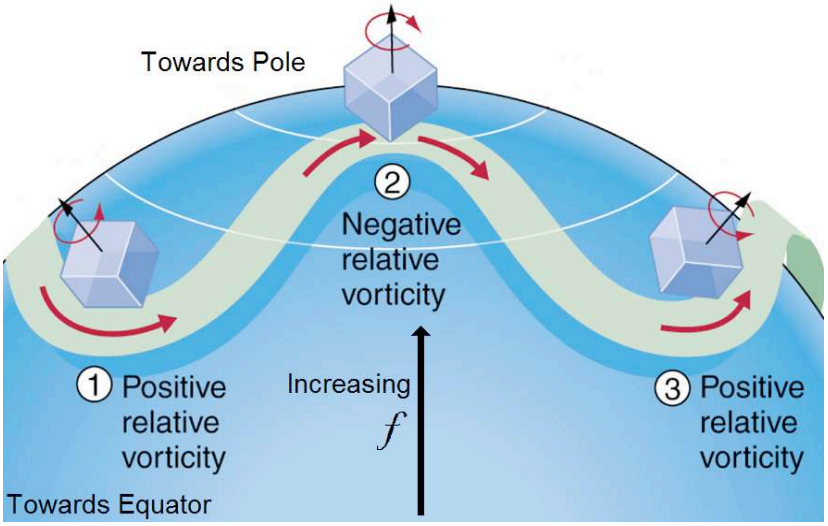


Figure II.14: The relationship between planetary vorticity and relative vorticity, when absolute vorticity is conserved. For an air parcel travelling in an eastward jet, in the northern hemisphere, which has an initial southward displacment. This can excite a Rossby wave which is stationary or westward relative to the flow.

Some of the largest amplitude planetary waves are the quasi-stationary mid-latitude Rossby normal modes. These are waves that correspond to a natural mode of oscillation of the Earth's atmosphere. Commonly observed in the middle atmosphere are oscillations with periods near to 2, 5 and 16 days, with zonal wavenumbers 3, 1 and 1 respectively.

In the stratosphere and lower mesosphere, these waves tend to maximise in the winter hemisphere due to the favourable propagation in the zonal mean winds, which are generally eastward. However, the quasi-two-day planetary wave has also been observed in the summertime at mid-latitudes in the upper mesosphere and lower thermosphere (e.g., Muller and Nelson, 1978). The zonal mean winds in summer prevent the propagation of these waves from source in the lower atmosphere. Therefore, the source of these waves has been suggested to be baroclinic instabilities in the summertime mesospheric jets (Plumb, 1983).

In the stratospheric polar regions planetary waves are very important because of their dissipation effects. Breaking of planetary waves produces meridional mixing of air in the polar vortex with warm air from lower latitudes and can even lead to the break up of the polar vortex. This effect prevents the winter polar stratosphere from becoming as cold as for the radiative equilibrium case.

II.4.4 Tides

Atmospheric Tides are global-scale waves, excited in two very different ways by the gravitational forces of the Moon and Sun and by the thermal actions of the Sun. Early observations connecting ocean tides with the Moon's transit across the meridian were made as far back as 320 B.C., prompting the idea that the Moon exerted an attraction on the water (Chapman and Lindzen, 1970).

Newton realised that these tidal forces must also affect the atmosphere as well as the oceans, but assumed that it would be too small for detection. Laplace's theoretical calculations indicated that it would be possible to directly measure a lunar tidal signal in pressure changes of ½ mm in mercury at ground level. Unfortunately, this is still a small change to observed and required a very large data set to extract the signal. This estimate of ½ mm in mercury, however, turned out to be an overestimate by about a factor of four. Laplace also knew of a daily variation, with a pronounced semi-diurnal component, which was much larger than his estimate for the lunar tide. He thought that these might be thermally excited but did not study this in any detail (Chapman and Lindzen, 1970).

Atmospheric tides can be now characterised by their sources. The Moon's gravity forces the lunar atmospheric tide, while solar atmospheric tides can be excited by either the absorption of solar radiation or the gravitational forcing from the Sun. For the solar tides, a diurnal variation in solar heating, where radiation is absorbed in the Earth's atmosphere, excites the tides. As the radiation enters the atmosphere, from above 150 km atomic oxygen is the most abundant atmospheric constituent and this absorbs the short wavelength EUV (extreme ultraviolet) solar radiation. Increasingly longer wavelengths are absorbed as the solar radiation approaches the Earth's surface. Molecular oxygen, O_2, absorbs the far-ultraviolet (FUV) radiation at altitudes near 100 – 150 km and ozone, O_3, absorbs the ultraviolet (UV) radiation between about 30 and 70 km. Infra-red (IR) is absorbed by water vapour (H_2O) in the lowest part of the atmosphere at tropospheric altitudes. Figure II.15a shows the heights at which tidal heating is excited and the main absorbing species responsible for this. Figure

II.15b shows the distribution of heating with latitude and Figure II.15c shows the distribution of heating with local time, where SR = sunrise, N = noon and SS = sunset. Notice that Figure II.15c has almost a square wave type pattern (Forbes, 1995).

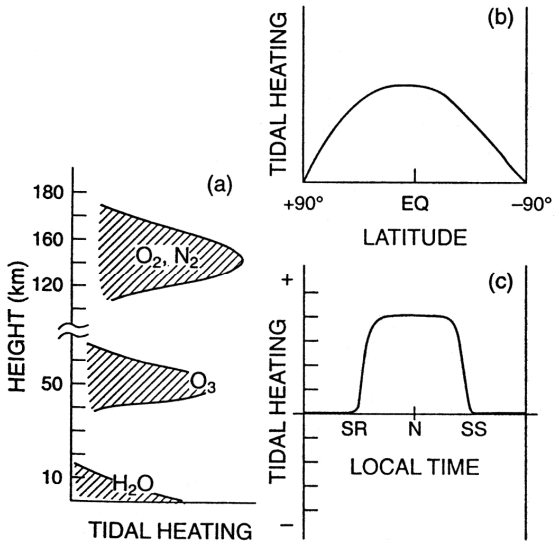


Figure II.15a, b and c: The variation in tidal heating with a) height including the major chemical species which contribute to the heating, b) with latitude and c) with local time, SR=sunrise, N=noon and SS=sunset (After Forbes, 1995).

This square wave of heating acts to produce tidal oscillations which have periods that are an integer fraction of a solar day (i.e., harmonics of the solar day). S_1 and S_2 refer respectively to the solar diurnal and semi-diurnal tides which have periods of 24 and 12 hours, respectively. Their lunar counterparts are the M_1 and M_2 tides.

Tides are a special class of buoyancy wave. The restoring force acting on atmospheric tides is gravity. Tides are affected by the Earth's rotation because of their large horizontal scales. They are also observed in all types of atmospheric fields including: wind, temperature, pressure, density, and geopotential height. Even though there is little, if any, tidal forcing due to solar heating in the upper mesosphere and lower thermosphere (~ 80 – 100 km), measurements of winds and temperatures exhibit strong tidal signatures in this region, due to the vertical propagation of the tides (Andrews *et al.*, 1987).

Lunar atmospheric tides are only about 5 – 10% as big as the solar tides. Unlike the solar tides, the lunar atmospheric tides are entirely driven by gravitational forces. The lunar gravitational acceleration decreases proportionally to the inverse square of the distance from the centre of the Moon. This acceleration is therefore not uniform around the Earth, depending on the Moon's location. Atmospheric air parcels at various locations on the Earth therefore experience slightly different forces. As the Earth

rotates during a lunar day (24.8412 h on average) an air parcel at the equator successively passes twice through regions of westward and eastward acceleration, comprising two lunar semi-diurnal cycles of period 12.4206 h, this corresponds to the Lunar M_2 tide (a more detailed study of the lunar tide is given in Chapter V). There is also a monthly periodicity to the forcing as the Moon cycles between the Northern and Southern Hemispheres of the Earth. In addition to the direct forcing of the Moon's gravity on the atmosphere, lunar atmospheric tides are indirectly forced by the Moon's gravity through the deformation of the Earth's surface and the oceans.

There are also two types of tidal behaviour; migrating tides, which refer to tides that follow the motion of the Sun or Moon around the Earth, in the case of the solar or lunar tides respectively. There are also non-migrating tides which have the same forcing mechanism and periods as migrating tide (i.e. the Sun or Moon), but which do not follow the motion of the Sun or Moon (i.e., departures from the migrating tides). Non-migrating tides are forced, for example, by the topology of the land, fixed geographic heat sources or non-linear coupling with other waves.

To further investigate tidal theory, we will look in depth at the solar migrating tides. The solar heating is only active during the hours of daylight. An example of this variation, in local time, is given in Figure II.15c. A Fourier analysis of this curve includes a steady-state mean component, a diurnal (24-hour) component, a smaller semi-diurnal (12-hour) component, and even smaller ter-diurnal (8-hour) component and so on. The response of the atmosphere to the heating can also be decomposed into a mean steady state, plus diurnal, semi-diurnal, ter-diurnal and higher frequency components.

In 1882, Kelvin raised the question as to why the semi-diurnal tide, at the surface seen in pressure observations, was larger and more regular than the diurnal tide. It is now known that the semi-diurnal tide has a larger vertical wavelength (> 100 km) than the diurnal tide. This allows it to be efficiently excited by deep ozone heating which is present near the stratopause. The excitement by ozone heating also makes the semi-diurnal tide regular, as the ozone is uniformly distributed in the stratosphere and its heating is very regular. The diurnal tide, however, has a more complicated behaviour. Between 30°N and 30°S it is able to propagate vertically with a shorter vertical wavelength of ~ 28 km. Poleward of 30° it is trapped near to its forcing region. The water vapour found in the troposphere is an effective exciter of the diurnal tide; however, this heating is intermittent in both space and time (Andrews *et al.*, 1987).

II.4.4.1 Classic Tidal Theory

Atmospheric tides require a further level of complexity compared with the gravity waves and Rossby waves. This is given in the form of Laplace's tidal equation. Laplace's tidal equation is an Eigen-value problem from which the meridional structure of the tide can be determined from its solution. The equation can be derived from the primitive equations (e.g., Chapman and Lindzen, 1970; Andrews *et al.*, 1987; Randall, 2000).

The meridional structure of tides is given by Hough functions, $\Theta_n(\varphi)$ which are complete between the latitudes of $-90 \le \varphi \le 90$. The subscript n denotes the possibility of multiple solutions. The meridional structure equation or the Laplace Tidal Equation is given by:

$$\frac{d}{d\mu}\left[\left(\frac{1-\mu^2}{\nu^2-\mu^2}\right)\frac{d\Theta_n}{d\mu}\right]-\frac{1}{\left(\nu^2-\mu^2\right)}\left[\frac{s}{\nu}\left(\frac{\nu^2+\mu^2}{\nu^2-\mu^2}\right)+\frac{s^2}{1-\mu^2}\right]\Theta_n+\varepsilon_n\Theta_n=0 \qquad \text{II.20}$$

Where $\nu \equiv \sigma/2\Omega$ is the normalised frequency and $\mu = \sin\varphi$ a latitude variable. The zonal wave number, s, represents the number of wave crests that occur along a latitude circle, while the sign of the wave number indicates the direction of propagation, positive wave numbers denote westward propagation and negative denote eastward.

Also ε_n is a dimension less quantity, know as a separation constant, which is usually described in terms of h_n, the equivalent depth, it is given by:

$$\varepsilon_n = \frac{4\Omega^2 a^2}{g h_n} \qquad \text{II.21}$$

The vertical structure equation is given by:

$$\frac{d^2 W_n}{dp^2}+\frac{S_p}{g h_n}W_n=-\frac{R}{g h_n c_p}\left(\frac{J_n}{p}\right) \qquad \text{II.22}$$

Where, S_p is the static stability, which depends on pressure, p (i.e., height), c_p the specific heat capacity at a constant pressure and R the ideal gas constant. J is a perturbation in heating and W is a perturbation in vertical velocity for a particular Hough mode. These two components are summed to give the heating rate, Q, (e.g., due to absorption of solar radiation) and the vertical component to the velocity, w, which are both sums over all possible solutions giving:

$$w = \sum_n W_n(p)\Theta_n(\varphi) \text{ and } Q = \sum_n J_n(p)\Theta_n(\varphi) \qquad \text{II.23}$$

Tides are forced oscillations, in which the frequency and zonal wavenumber of the forcing are specified. This is in contrast to free oscillations in which the frequency and zonal wavenumber are given by the solution of the eigenvalue problem. For a given frequency σ and wavenumber s, the Laplace Tidal Equation (II.20) will yield a generally infinite set of equivalent depths, h_n, (the eigenvalues) and Hough functions, $\Theta_n^{\sigma,s}$, (the eigenfunctions). The n denotes the meridional index, which gives information on the number of latitudinal nodes. Even n numbers correspond to Hough modes which are symmetric about the equator and Odd n numbers correspond to anti-symmetric modes. Hough modes can also be written as $\Theta(L,M)$ where there are $M = |n| - s$ nodes of the Hough function between the poles, excluding the polar nodes and L refers to the zonal wavenumber. Negative meridional indices indicate that the mode is evanescent and will not propagate vertically (c.f. the semi-diurnal tide being larger than the diurnal tide as a number of diurnal modes have negative meridional index).

Each Hough mode is a function of latitude and may be expressed as an infinite sum of associated Legendre polynomials (solutions to differential equations). In basic terms, the Hough functions describe latitudinal structure of the waves (Andrews *et al.*, 1987).

As an example, taking the migrating solar semi-diurnal tide, $s = 2$ and $\sigma = 2$ taking only the symmetric modes, Figure II.16 shows the latitudinal distribution for the first three Hough functions. The vertical structure of each eigensolution is also determined with the vertical structure equation II.22.

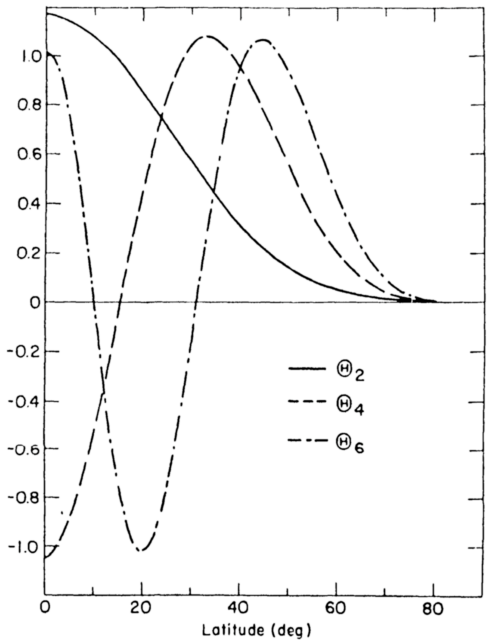


Figure II.16: The latitudinal distribution of the first three symmetric Hough functions for the migrating solar semi-diurnal tide (After Chapman and Lindzen, 1970).

The Global Scale Wave Model (GSWM) uses these Hough functions to calculate the predicted amplitudes of tides in the atmosphere, which are estimated to maximise in the MLT region (Hagan and Forbes, 2003 and references therein).

Note that tides exhibit downward phase propagation with the upward transport of wave energy and that in the MLT region there is little forcing from solar heating which implies that the tides have been created elsewhere and propagated vertically. The general mathematical expression for the magnitude of a tide in the atmosphere at a given time and longitude is given by:

$$A\cos\left(\sigma t - s\lambda - \phi\right) \qquad \text{II.24}$$

Where in this case ϕ is the phase of the tide and A is the amplitude

II.5 Summary

This chapter introduced the dynamics of the atmosphere, including the middle atmosphere, whose dynamics are dominated by winds, waves and tides. The importance of the dynamics in this region in terms of general circulation (i.e., the Dobson Brewer circulation) has been discussed. The dynamics in the middle atmosphere have also been presented and shown to include gravity waves, planetary waves and tides. The dispersion relations of some of these waves have been

discussed, which shows the conditions under which these waves are able to propagate. Classical tidal theory has also been discussed from which it is possible to determine the structure of forced waves in the atmosphere. The forcing is from gravity and solar heating in the case of the solar driven tides and purely gravity in the case of the lunar driven tides.

In conclusion, the dynamics of the middle atmosphere are vitally important in understanding the circulation of the entire atmosphere. There are still phenomena occurring in the middle atmosphere which are as yet not well understood.

Chapter III

III. Remote Sensing of the Stratosphere, Mesosphere and Lower Thermosphere

A wide variety of remote and *in situ* observational techniques have been used to investigate the waves of the middle atmosphere. These techniques have revealed waves with horizontal wavelengths ranging from a few km to global scale and periods ranging from a few minutes to over 2 weeks.

This chapter details two of the techniques used in this thesis. In particular, focus is given to the derivation of the MLT winds from a meteor wind radar and geopotential height measurements of the troposphere, stratosphere and mesosphere using a satellite-borne microwave limb sounder.

III.1 Meteor Wind Radar

Meteor wind radars measure radio waves reflected from the plasma trail left by the ablation of a meteor at heights of about ~ 80 -100 km. These trails are carried by the winds in the mesopause region. Therefore, using the Doppler shifts of reflected radio waves, the velocities of the winds can be calculated. Because modern meteor radars can operate continuously and automatically they can be more effective than other techniques, which require constant maintenance. Unlike optical techniques they are also not limited by weather conditions or by sunlit hours.

III.1.1 Meteors

Meteors, sometimes know as "shooting stars" are bright bursts of light, sometimes the brightest objects seen in the night sky. These flashes of light are caused by the de-excitation of material from an ablating meteoroid in the Earth's atmosphere.

Meteoroids are small chunks of interplanetary debris, possibly left behind by passing comets or asteroids. Meteoroids are sometimes captured by the Earth's gravitational field and as they travel through the atmosphere they are heated by friction. When its surface temperature reaches ~ 2,200 K (this usually occurs between about 80 and 90 km) the meteoroid's surface begins to sublimate. This produces hot excited vapours, which are left behind as a trail of ionised atoms. The trail formed is approximately cylindrical and stretching out for 10 – 15 km behind the meteoroid. When these stripped atoms recapture electrons, photons are emitted, giving the bright trail of light seen following the meteoroid.

The speed of meteors typically ranges from approximately 11 - 73 kms^{-1}. This minimum speed is due to the Earth's gravity. The maximum speed of a meteoroid comes from the particle's motion around the Sun at the distance of the Earth (1 AU), which has a maximum of 42.5 kms^{-1} (i.e., the Sun's escape velocity). Anything faster than this would not be orbiting the Sun and would therefore probably be passing through the solar system, which is rare. So, anything moving in a fast retrograde orbit of the Sun compared to the Earth would have its own velocity plus that of the Earth, 30.3 kms^{-1} at perihelion (i.e., giving a theoretical maximum of ~ 72.8 kms^{-1}). High speed meteors have been observed in the Leonids meteor shower, which follow a retrograde orbit of the Sun (Ceplecha *et al.,* 1998).

Occasionally, intense visual meteor displays fill the sky with hundreds of meteor trails. These displays are called meteor showers and many can be predicted, as they repeat every year when the Earth passes through the path of a comet. Each day as many as 4 billion meteors fall to the Earth, with their total mass being several tones (Ceplecha *et al*., 1998).

III.1.1.1 The Meteor Trail

Immediately after the ionised trail is formed by the ablating meteoroid, it begins to dissipate, by recombination, ambipolar diffusion and atmospheric turbulence. These all cause a decrease in the trail's electron density. The most important of these effects is ambipolar diffusion, where the charged atoms and free electrons disperse away from the original trail, taking only a few seconds to disperse. If the meteor trail is vaguely aligned with the Earth's magnetic field lines it can have the effect of prolonging the lifetime of the trail, although this effect is less significant below about 95 km (Ceplecha *et al*., 1998).

The meteor trail is also advected by the neutral wind. By observing how the meteor trail drifts with time, deductions can be made about the speed and direction of the atmospheric wind at heights in which meteors are observed.

III.1.1.2 Radar Reflection from Meteor Trails

It is possible to use radar (**Ra**dio **D**etection **a**nd **R**anging) to detect meteor trails in the atmosphere. The large mass of the positive ions in the meteor trail prevents them from being significantly oscillated in the presence of the radio wave, but the free electrons have a small enough mass to reflect the radio signal through Thompson scattering. The scattering from individual electrons is randomly phased, but if the electrons are aligned (in a meteor trail for example), the scattering can be strongly coherent (McKinley 1961), and give a strong enough reflection to be observed without the need for the high energy transmissions.

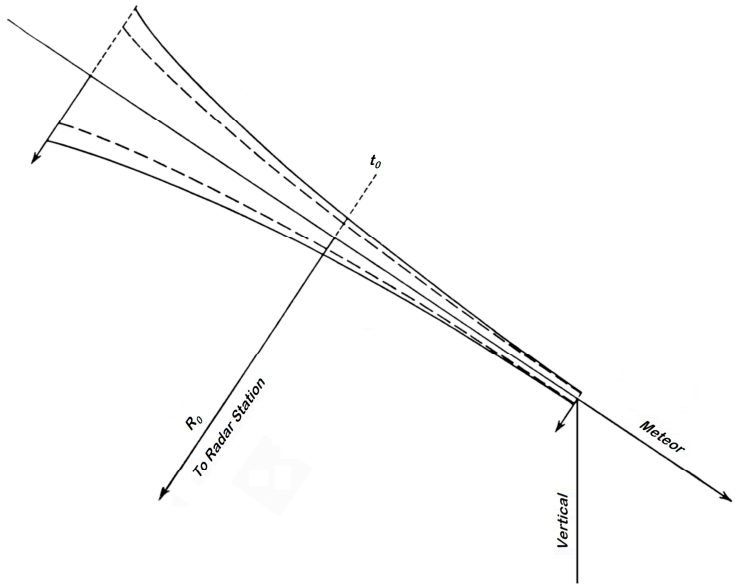


Figure III.1 - The Meteor Trail, an approximately cylindrical trail left behind by an ablating meteor, which diffuses radially with time (After McKinley 1961).

Only a small fraction of the incident energy transmitted by a radar is reflected back to the ground and these signals can be detected by receiver antennas. The reflection of radio waves off meteor trails is specular reflection (similar to reflection from a mirror). Therefore for the case of backscatter (transmitter and receiver at the same site), the trail must be oriented orthogonally to the radio wave to reflect the wave. Trails which are not orientated orthogonally will reflect the signal away from the radar. This, therefore, limits the number of meteors which can be seen by a radar.

Radar allows detection of meteors below the visual limit. The majority of meteor trails detected are produced by meteoroids of mass smaller than 10^{-5} kg (a radius of ~ 0.1 cm). The sensitivity of the radar is dependant on the power of the transmitted pulse and the antenna parameters e.g., the radiation pattern.

The way in which a radio wave reflects from the meteor trail depends upon the density of free electrons in the meteor trail, as reflection is accomplished by scattering of the radio wave by the individual electrons in the trail. There are essentially two types of reflection; those from trails with low electron line densities ($\leq 10^{14}$ electrons per meter), "underdense" meteor trails and those of high electron densities ($\geq 10^{14}$ electrons per meter), "overdense" meteor trails. The transition from one form to the other is not sharply defined (Ceplecha *et al.*, 1998). There is a third type of radio reflection, which is know as a "head echo", in which radio waves reflect from a cloud of ionisation at the head of the meteor. This type of reflection is not observed with the small transmitter powers used by meteor wind radar as the reflections are much less coherent, but it is possible using the much higher transmitter powers of, for example, Incoherent Scatter radar.

An example of the signal received from an underdense meteor echo is shown in Figure III.2. It consists of a very steep rise in echo power, typically in a few hundredths of a second. Once the trail begins to diffuse, loss in coherence of the scattered waves causes an exponential signal decrease in the echo power. The signal from underdense meteors lasts no more than a few tenths of a second. Underdense meteors cover a meteoroid mass range which is not observable by the naked eye. They are numerous and they also strictly obey the specularity rules. This is useful for calculating the position of the meteor reflection point in the sky and the direction of the meteor path in the atmosphere. The characteristic shape of an underdense meteor echo is very useful in distinguishing the meteor from other targets such as aircraft, interference or ground clutter.

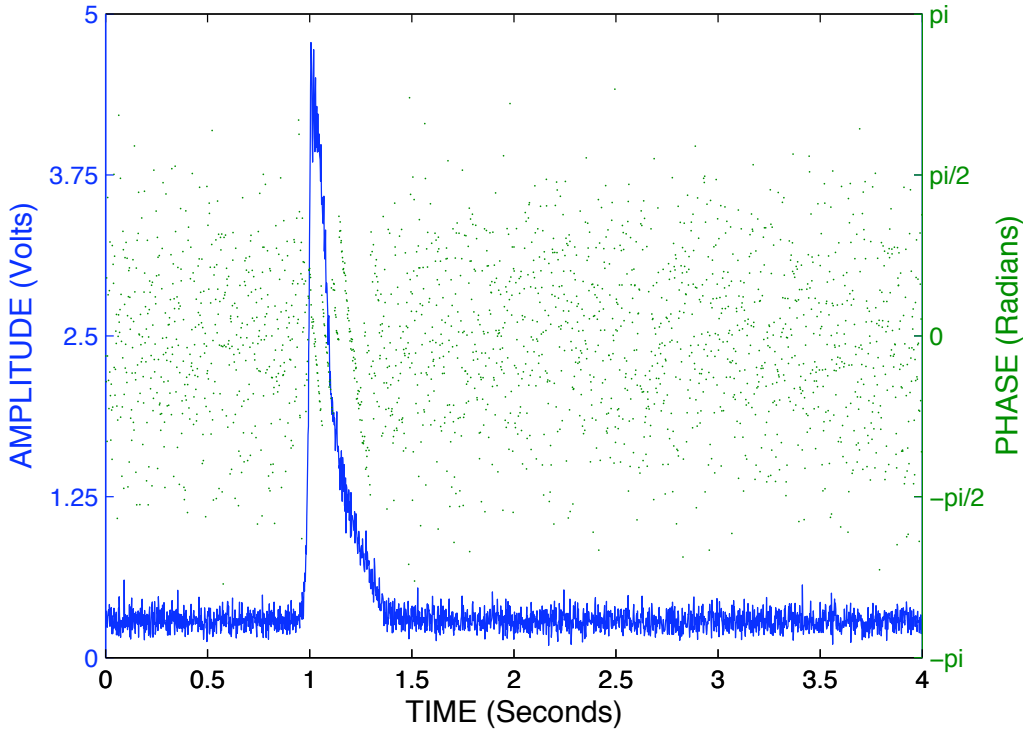


Figure III.2 - A typical underdense meteor echo, observed by the Esrange meteor radar. The blue line indicates the amplitude of the received power and the green dots indicate the phase. Note the sharp rise in amplitude when the meteor is detected and the consistency in phase at the same time.

Overdense meteors are less numerous. The reflections can last for a few seconds, during which the trail is often subject to severe distortions and shears by the strong winds. This leads to the received signal being irregular and unpredictable. Interference can occur in the received signal from multiple reflection points on different parts of the trail. As these reflection points are not stationary due to the wind, fast oscillations can be observed in the received signal power. Figure III.3 shows a typical overdense meteor observation. The irregular behaviour of the echo power in an overdense echo makes it much more difficult to distinguish from noise/interference, than is the case for an underdense echo. Because of this, many meteor radars record only underdense echoes so as to be confident of identifying only meteor scatter.

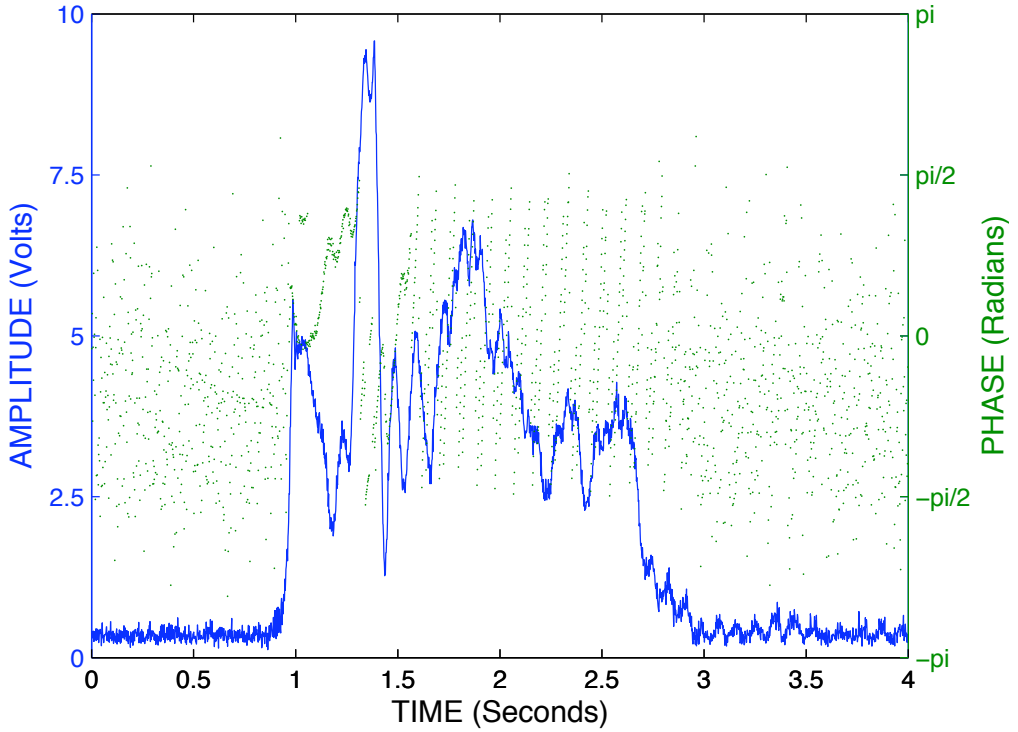


Figure III.3 - A typical overdense meteor echo, observed by the Esrange meteor radar. The blue line indicates the amplitude of the received power and the green dots indicate the phase. Note the oscillations in the signal are likely caused by the interference between signals reflected on different parts of the trail.

For the case of underdense echoes, if the initial radius of the meteor train is small compared to the wavelength of the transmitted signal, then the received signal is coherent. However, as the trail diffuses the radius can approach the size of the wavelength. If this happens, destructive interference can occur between the near and far sides of the trail and the returned signal is greatly attenuated. This gives a ceiling height for meteor trails which the radar of specific wavelength can measure, as the initial radius of a meteor trail is generally larger at higher altitudes.

III.1.2 Radio Scatter from Meteors and the Radar Equation

This section describes how the intensity of a signal scattered from a meteor trail can be estimated using the radar equation. The description will follow that of McKinley (1961) and Ceplecha *et al*. (1998). For an underdense meteor trail each electron may be regarded as an independent scatterer. The transmitted radio wave is able to penetrate the trail with little attenuation and each electron acts as a small Hertzian dipole in reflecting the radio wave. The total scattered radiation is thus the sum of the contributions from each individual electron.

The scattering cross section of a free electron is given by:

$$\sigma_e = \frac{4\pi e^4}{m^2 c^2} \sin^2 \gamma = 4\pi r_e^2 \sin^2 \gamma \qquad \text{III.1}$$

Where

e is the electron charge (1.60×10^{-19} C)

m is the electron mass (9.11×10^{-31} kg)

r_e is the classical electron radius (2.82×10^{-15} m)

c is the speed of light (3.00×10^{8} ms^{-1})

γ the angle between the incident electric field vector and the direction of the receiver.

For the case of backscatter $\gamma = 90^\circ$ and therefore $\sigma_e \simeq 1.0 \times 10^{-28}$ m^2.

The power flux transmitted by the radar at the distance of the trail is dependant on the initial transmission power, P_T, the gain of the transmission antenna, G_T, and the distance to the trail, R. The incident power flux, Φ_i, is then given by the equation:

$$\Phi_i = \frac{P_T G_T}{4\pi R^2} \qquad \text{III.2}$$

The power flux, Φ_e, which is then returned to the receiver site due to the scatter from one electron is given by the equations:

$$\Phi_e = \Phi_i \frac{\sigma_e}{4\pi R^2} \qquad \text{III.3}$$

And the effective collecting area, A_{eff}, of the receiver antenna by:

$$A_{eff} = \frac{G_R \lambda^2}{4\pi} \qquad \text{III.4}$$

Where G_R is the receiver gain, and λ the wavelength.

By combining equations III.3 and III.4, the received power, P_R, of the scatter from a single electron can be shown to be:

$$P_R = \frac{P_T G_T}{4\pi R^2} \cdot \frac{\sigma_e}{4\pi R^2} \cdot \frac{G_R \lambda^2}{4\pi} \equiv \frac{P_T G_T G_R \lambda^2 \sigma_e}{64\pi^3 R^4} \qquad \text{III.5}$$

Where P_T is the transmitted power, G_T and G_R are the gains of the transmitter and receiver antennas respectively, R is the distance between the transmitter/receiver and

the electron in the meteor trail (assuming the transmitter and receiver are co-located) and λ is the radio wavelength.

To now consider a meteor trail consisting of many electrons rather than just one, we first assume that the trail diameter is small compared to the incident wavelength. The electrons are assumed to be evenly distributed in a line on the axis of the trail. Thus for an element of the line, all the electrons should scatter in phase.

The electric field, (e.g., E_e) rather than power fluxes (e.g., Φ_e) are added to find the resultant field of the trail. The maximum amplitude of the field is given by:

$$\left(E_R\right)_0 = \left(R_0\,\lambda/4\right)^{1/2}\left(2Z_0\Phi_e\right)^{1/2} q\left(C^2+S^2\right)^{1/2} \qquad \text{III.6}$$

Where Z_0 is the wave impedance of free space, R_0 is the orthogonal distance to the trail and C and S are the Fresnel integrals of optical diffraction theory. With q electrons per metre along the trail, the power flux can then be given by:

$$\Phi_R = \left(E_R\right)_0^2 \big/ 2Z_0 = \Phi_e\left(R_0\,\lambda/4\right) q^2\left(C^2+S^2\right) \qquad \text{III.7}$$

Thus the power delivered to the receiver is given by:

$$P_R = \Phi_R A_{eff} = \Phi_e\left(G_R\lambda^3 R_0/16\pi\right) q^2\left(C^2+S^2\right) = \frac{P_T G_T G_R \lambda^3 q^2 r_e^2}{64\pi^2 R_0^3}\left(C^2+S^2\right) \qquad \text{III.8}$$

If there are several Fresnel zones either side of the orthogonal point then $C^2+S^2 \approx 2$ and contributions from remoter parts of the trail are negligible. The ability of a radar to detect the underdense meteor trail then depends greatly on the initial radius of the meteor trail, r_0. This is defined as the distance from the meteor axis to the point where the ablated atoms slow to thermal speeds, due of collisions with the neutral atmosphere. The initial radius is a function of the meteor height and its velocity.

If r_0 is comparable to or greater than the transmitted wavelength, the backscatter from the electrons in the near and far parts of the trail will not be in phase and these scattered components will destructively interfere. If $r_0 \ll \lambda$, then on the initial formation of an underdense meteor trail, a radio-wave reflection from the trail will have the initial power:

$$P_R(0) \equiv \frac{P_T G_T G_R \lambda^3 q^2 r_e^2}{32\pi^2 R^3} \qquad \text{III.9}$$

However, the finite radius of the actual trail, when first formed, reduces the initial echo amplitude depending on the transmitted wavelength, λ. This attenuation factor is given in the equation:

$$\sigma_r = \exp - \left(\frac{8\pi^2 r_0^2}{\lambda^2} \right) \qquad \text{III.10}$$

Also, the finite velocity of the actual trail, when first formed, reduces the initial echo amplitude, as the ambipolar diffusion of the trail is dependant on height. This attenuation factor is given in the equation:

$$\sigma_v = \exp - \left(32\pi^2 \frac{D}{V} \left(\frac{R}{2\lambda^3} \right)^{1/2} \right) \qquad \text{III.11}$$

Where D is the average ambipolar diffusion coefficient and V is the velocity of the meteor. Therefore the returned power at time, t, becomes:

$$P_R(t) = \frac{P_T G_T G_R \lambda^3 q^2 r_e^2 \sigma_r \sigma_v}{32\pi^2 R^3} \qquad \text{III.12}$$

Both the σ_r and σ_v attenuation factors vary with wavelength. This effectively gives a ceiling to the meteor detection, depending on the transmitted wavelength, for example the Esrange SKiYMET meteor wind radar operates at a transmission frequency of 32.5 MHz and so has a wavelength of ~ 9.22 m. This radar has an underdense echo ceiling between 100 and 110 km. The ceiling is not a precise height as it varies depending on the meteor's velocity and entry angle.

After the trail's initial formation it undergoes dispersion, mainly due to ambipolar diffusion. This expands the trail radially and causes the echo power to decay exponentially with time as the signal becomes less coherent. The power received, P_R, at a time t after the trail formation is given by:

$$P_R(t) = P_R(0) \exp\left(-\frac{2t}{\tau} \right) \qquad \text{III.13}$$

Where the echo amplitude decay time, τ, is defined as the time take for P_R to decay to a factor of e^{-2} and is given by:

$$\tau = \frac{\lambda^2}{16\pi^2 D} \qquad \text{III.14}$$

This exponential decay can be seen in the underdense meteor echo of Figure III.2 after the sharp rise, the amplitude decays away exponentially.

III.1.2.1 Measuring Meteor Parameters

Radars with an interferometer based receiver array, which have multiple receiver antenna, can determine many parameters from the meteor echoes by comparing the phase information from each of the spaced-out receivers. This includes azimuth, elevation and height of the meteor; from which the angle of arrival can also be calculated. These can also enable calculation of other parameters such as the meteor radiant, entry speed and orbits. An example of an interferometer type of radar is the SKiYMET system, which is described later. There are also non-interferometric techniques, such as the UK beam system, also described later. The non-interferometry based radars only use the amplitude and time of the received signal and therefore height finding and precise location are not available.

Using the meteor trail as a tracer, neutral atmospheric wind can be determined from the radar echo. In calculating the winds, it is possible to use the method of "Doppler shift". After the formation of the meteor trails they often drift, due to the strong winds at this height. Radio signals reflected from these moving trails will be shifted in frequency, depending on whether the trail is moving towards the radar or away. The size of the shift can also be used to tell the speed with which the trail is moving with respect to the radar. It is there for possible to calculate a radial velocity for each meteor as shown by:

$$\Delta f = \frac{V_{Rad}}{\lambda} = \frac{V_{Rad} f}{c} \qquad \text{III.15}$$

Where,

Δf is the change in frequency of the received signal

V_{Rad} is the radial velocity,

λ is the transmitted wavelength,

f is the transmitted frequency and

c is the speed of light.

Using the phase information at each of the receivers in the interferometer system, the zenith angle of the meteor with respect to the radar site can also be obtained. Using the radial velocity and zenith angle, the line of sight horizontal velocity can then be calculated using:

$$V_{Horz} = V_{Rad} \sin(\vartheta) \qquad \text{III.16}$$

Where,

V_{Horz} is the horizontal velocity,

V_{Rad} is the radial velocity and

ϑ is the zenith angle.

In the case of a meteor radar with no interferometer, it is assumed that meteors fall at ~ 90 km and using the range information, from the time between pulses, the zenith angle can be found.

Figure III.4 shows the radial direction (away or towards the radar) and horizontal distance a meteor trail would move in one thousand seconds, displayed as red arrows. The data taken from the Ascension Island radar also shows the topography (mainly ocean) around Ascension Island, which is located at the centre of the Figure. If the speed of the horizontal winds is plotted against the azimuth of the meteor's trail, it is often possible to see a reasonable sine wave. This is because the winds from each meteor are generally flowing with the bulk motion. The analysis is presented in Figure III.5. Any departures from the sine wave probably indicate the presence of gravity waves. Therefore, by fitting a sine wave, to the meteor distribution in Figure III.5, the magnitude and direction of the bulk motion can be calculated. This is assuming a volume average and ignoring any small scale motions.

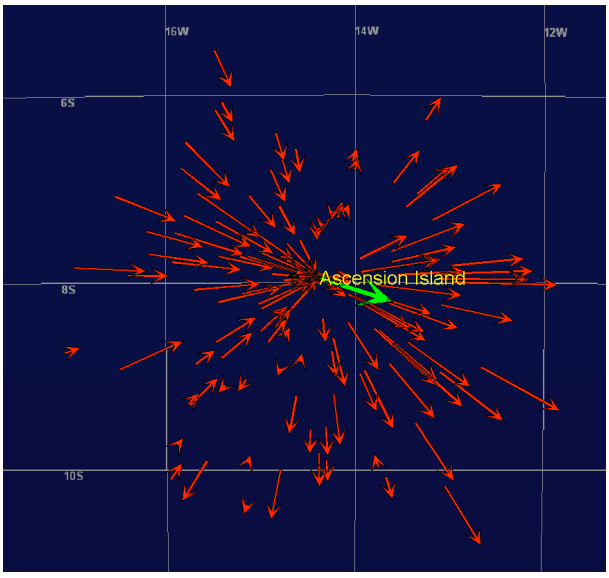


Figure III.4: Horizontal line of sight winds from the Ascension Island meteor radar on 28th January 2007 from 12:00 to 14:00 hours (UT), over the height 89 to 92 km. The red arrows indicate how a meteor trail will move over 1000 seconds calculated from the Doppler shift of each meteor and the Green arrow indicates the estimate of the bulk flow for all the meteors.

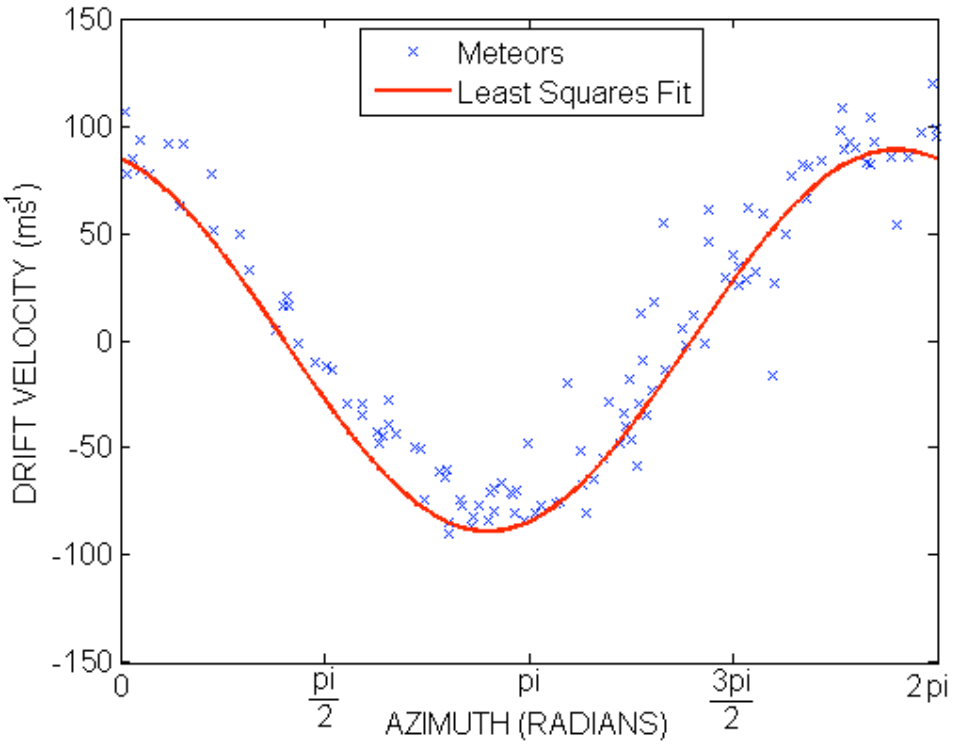


Figure III.5: Horizontal winds against azimuth. Each blue cross represents a single meteor event. Data is taken from the Ascension Island meteor radar on 14th November 2006 from 05:00 to 07:00 hours, over the height 89 to 92 km. The red line indicates the line of best fit using a least-squares sinusoidal fit.

Using this method assumes that only the influence of the neutral wind has a significant effect on the trail's behaviour. Other phenomena, such as trails which are closely aligned with the Earth's magnetic field or the fact that a trail's motions can vary in the presence of very strong electric fields (e.g., during high geomagnetic activity), are assumed not significant.

The meteor trails are also assumed to move with the bulk flow over the entire collecting area of the radar, up to ~ 300 km in radius; that there is also little change in the wind velocity over the 2 hour averaging period and that there is little wind shear over height range being measured. This poor horizontal resolution is not a problem when measuring the mean winds, tides and planetary waves (See Chapter II), as they all have much larger horizontal scales.

III.1.3 The SKiYMET Meteor Radar

The SKiYMET radar systems consist of one transmitter antenna and five receiver antennas. Figure III.6 shows a picture of one of the receiver antennas from the Rothera SKiYMET system. The radar transmits short pulses of energy in the form of very high frequency (VHF) radio waves from the transmitter antenna. Some of this

energy is scattered back from a meteor trail and is detected by the array of receiver antennas. The antennas with a three-element Yagi for the transmitter and five two-element Yagi for the receivers, have a broad radiation pattern designed to illuminate a large expanse of the sky. This is known as an "All-Sky" radar technique. Figures III.7a and b display the polar diagrams for the transmitter/receiver antennas, with maximum gain found at the four cardinal points of the compass and ~ 40 degrees elevation, with gain dropping off rapidly at angles below 30 degrees elevation (Hocking *et al.*, 2001).

Figure III.6: A SKiYMET meteor radar receiver antenna. Part of the Antarctic meteor radar located at Rothera (68°N, 21°E) station (Photograph by P. T. Younger).

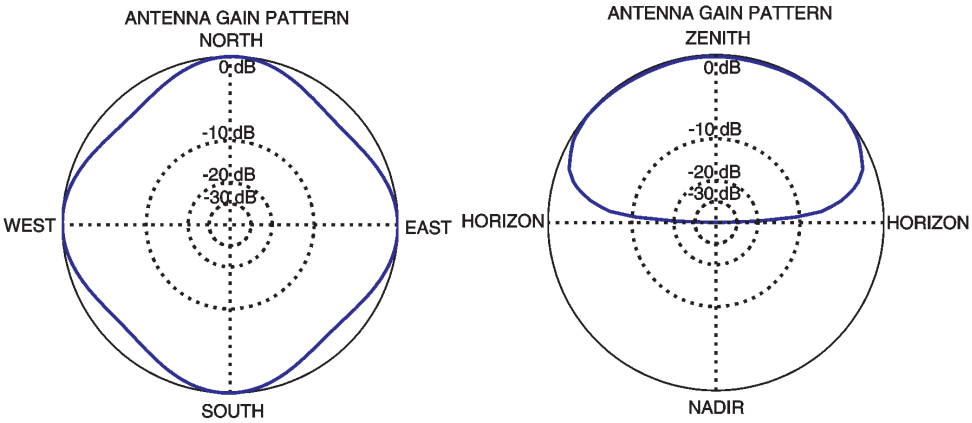


Figure III.7a, b: The polar diagram of the receiver / transmitter gains of the antenna used in a SKiYMET radar system, a) top-down (overhead) projection and b) elevation projection.

The receiver antenna array is arranged as an interferometer, spaced by 2.5 wavelengths and 2 wavelengths along two orthogonal axes. The phase difference between the signals arriving at each of the antennas is then used to determine an unambiguous zenith and azimuth angle of arrival. Then, together with the range information, the position of the meteor can be accurately located in the atmosphere.

The number of radio pulses transmitted per second, the pulse repetition frequency (PRF), establishes the maximum unambiguous range of the radar. However, as shown in Figure III.8, it is possible to get range aliasing, For example, as shown in Figure III.8 the second received signal could be from the first transmitted pulse reflecting in the meteor region or it could also be a return signal from the second pulse reflecting lower in the atmosphere. It would not be possible to distinguish which pulse was being reflected in the returned signal. It is possible to overcome this problem by using pulse coding, a technique used in incoherent scatter radars for example. However, using the *a-priori* knowledge that the meteor region is between ~ 70 and 110 km and reflections are likely to be coming from these heights, it is possible to estimate which pulse is being reflected in this region. However, this is not flawless as meteors seen at large horizontal ranges from the radar can have range aliasing between 70 km and 110 km, giving ambiguous detections (Kingsley and Quegan, 1992; Hocking *et al.*, 2001).

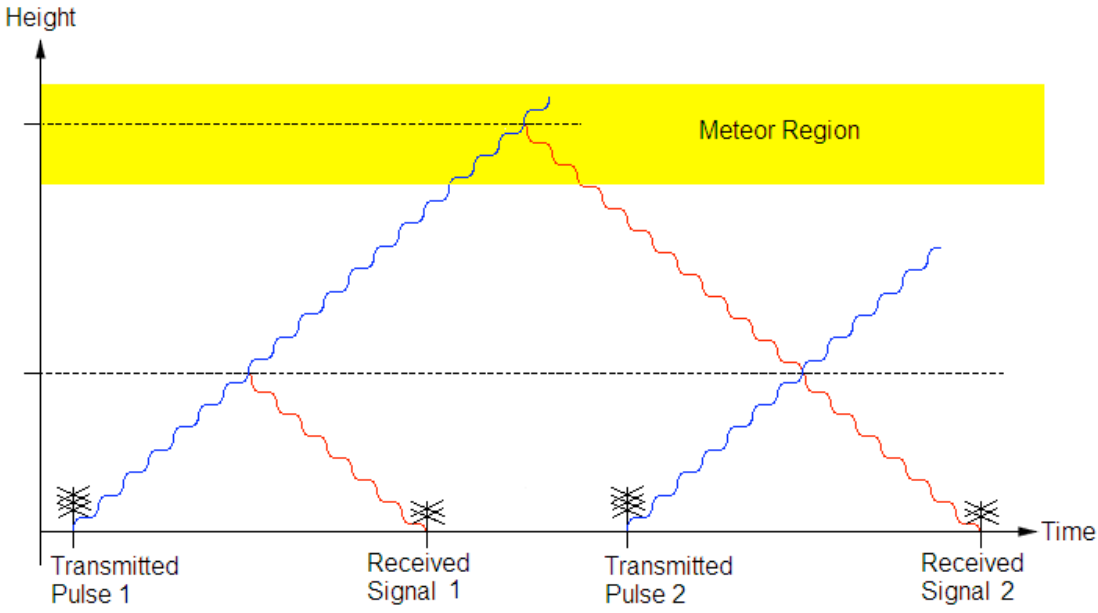


Figure III.8: An example of range aliasing, where the second received signal could have been transmitted from pulse 1 or pulse 2 and would give the same signal.

III.1.3.1 Meteor Echo Distributions

The SKiYMET systems have a height resolution of ~ 2 km and meteors are detected between 70 and 110 km in height. Anything found outside this is rejected. The curvature of the Earth's surface has also been taken into account in the heights of meteor detections. Figure III.9 displays a top-down (overhead) projection of the location of each meteor collected by the Esrange meteor radar from the date of its commission in 1999 to February 2007 in the narrow height band of 89 to 91 km. This contains over two million meteor detections. A single meteor event is represented by a red dot and the radar is located at the centre. From the figure, it is clear to see circles in which no meteors are detected. This is due to the receiver having to be switched off during the transmission of a pulse, to avoid saturating and damage to the sensitive receiving equipment. Also note that there are no meteor reflections collected directly over the radar. This is due to the angle which would be required for a reflection. These meteors would need to be travelling tangential to the Earth's surface for a reflection to occur. Therefore the meteors would have travelled through many hundreds of km of atmosphere, and most likely completely ablated. These reflections are therefore very rare and contain large errors, so are thus excluded.

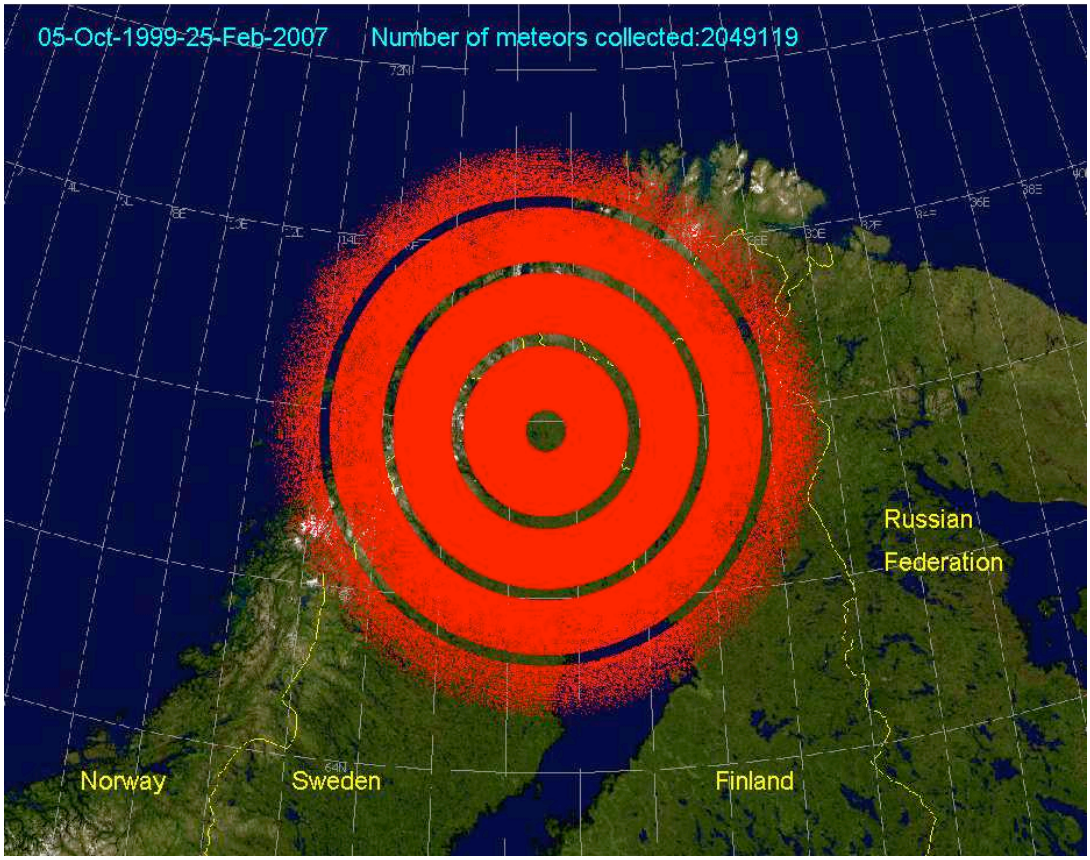


Figure III.9: The location of all meteors collected between 89 and 91 km by the Esrange meteor radar between October 1999 and February 2007. Each dot represents a single meteor event.

Figure III.10a, b and c display the distribution in time, height and range, respectively, of the meteor detections. Figure III.10a shows that there are many more meteor events during the morning sector of the day between 00:00 and 10:00 UT and many less in the evening, say, 14:00 to 22:00 UT. This is due to the fact that as the Earth revolves around the Sun, the leading edge will sweep up more sporadic meteors (Hocking *et al.*, 2001). This is analogous to bugs hitting the front windscreen of a car. Therefore when the station it is on the leading side of the Earth at dawn, we expect the detection of more meteors, and when it's on the trailing edge at dusk, we expect less. Figure III.10b displays a typical distribution of meteors with height for this particular radar and transmission frequency. It shows that the meteors are strongly peaked at ~ 90 km with a Gaussian distribution. Figure III.10c displays a typical distribution of meteors with range. It shows that the meteors are strongly peaked at ~ 150 km in range and the drop off in count rate is exponential after that (Hocking *et al.*, 2001).

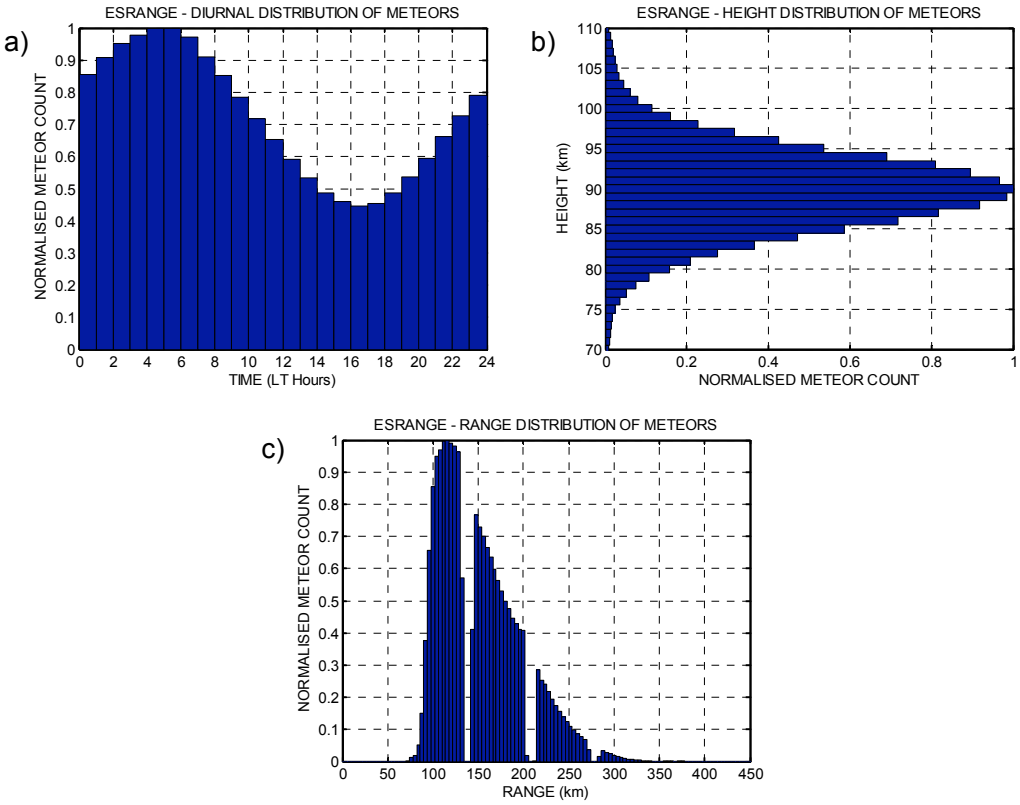


Figure III.10a, b and c: Meteor distributions produced by a SKiYMET radar for unambiguous meteors. Figures show the distributions over 9 years of data collected from the Esrange meteor radar. a) Count rate as a function of time of day over one day. b) Height distribution of meteors. c) Range distribution of meteors.

In a typical day a SKiYMET radar will detect around 4,000 meteors, but this can rise to as many as 8,000 meteors or as little as 1,000 depending on season and location of

the station. There are also many more meteors detected during meteor shower events. Figure III.11 displays the distribution of meteors received over Rothera (68°S) in a typical day, where each cross represents an individual meteor event and Figure III.12 displays the distribution of meteors collected in an average year from ~ 7 years of Esrange data, October 1999 – June 2007. There is a clear annual cycle in the count rates seen over each station each year. This cycle is believed to be astronomical as a result of the sources of sporadic meteors (Hocking *et al.*, 2001).

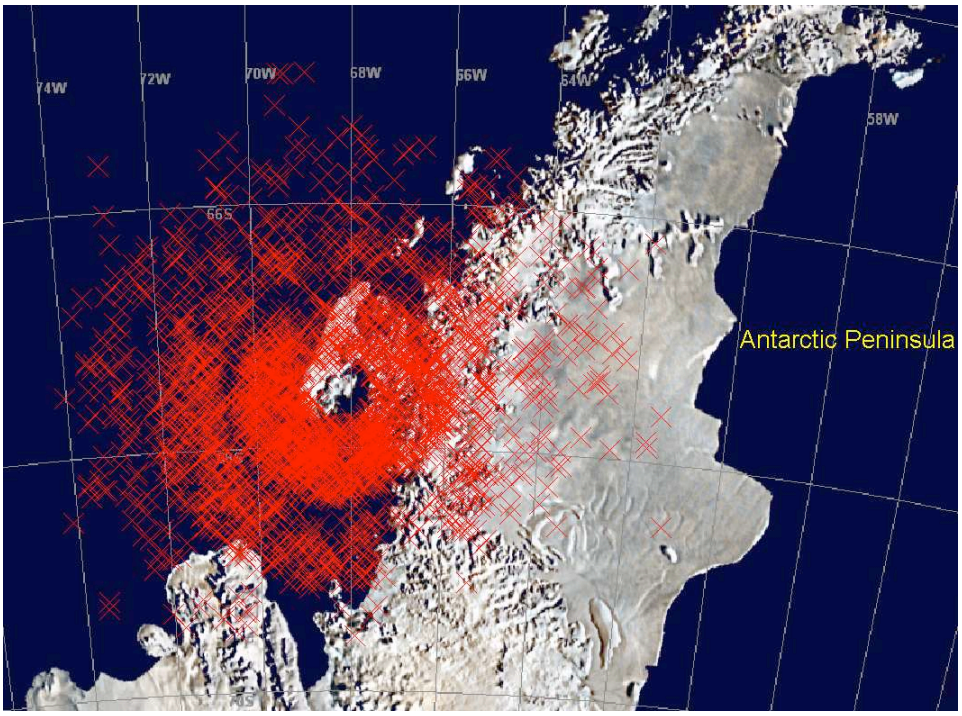


Figure III.11: The location of each meteor, ~ 3,000, collected between 78 and 110 km by the Rothera meteor radar on 2nd September 2008. Each cross represents a single meteor event.

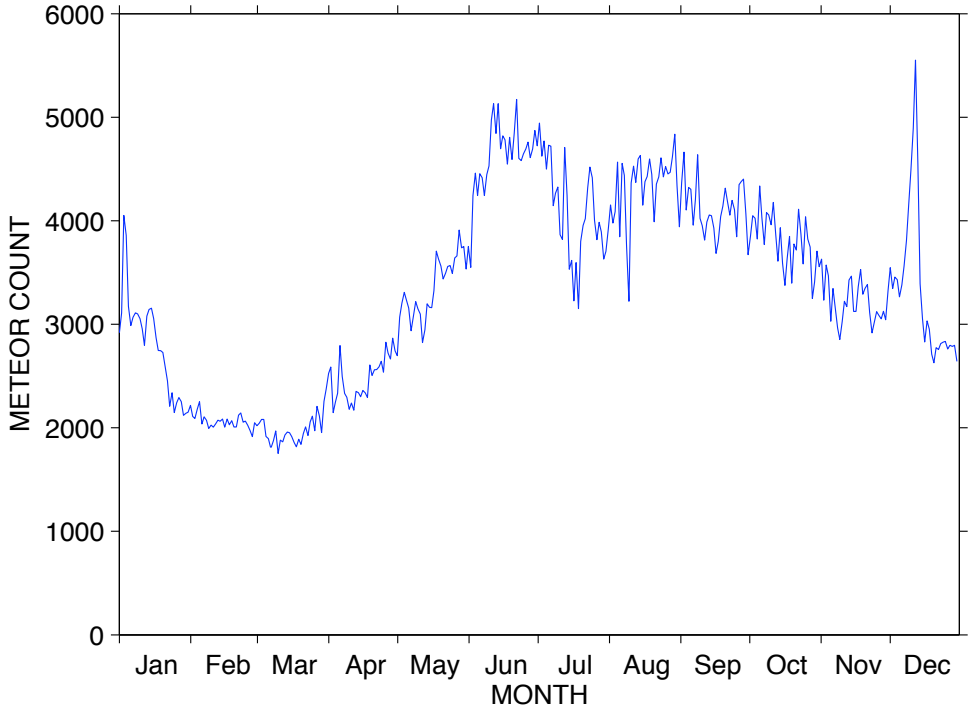


Figure III.12: The change in meteor count rate over a composite year from Esrange between 1999 and 2007. Also note the Geminids meteor shower spike in mid-December and the Quadrantids in early January.

Details of the meteor radars used in this study and owned by the University of Bath are given in Table III.1.

Table III.1, SKiYMET Radar information

Radar	**ESRANGE**	**ASCENSION**	**ROTHERA**
Location	67.9°N, 21.1°E	8.0°S, 14.4°W	67.5°S, 68.0°W
Pulse Repetition Frequency	2144 Hz	2144 Hz	2144 Hz
Transmission Frequency	32.5 MHz	43.5 MHz	32.5 MHz
Transmitter Power	6 kW	12 kW pre Oct 05 / 6 kW post Oct 05	6 kW
Time zone	GMT + 1	GMT	GMT - 3
Date of Commission	5th October 1999	26th May 2001	12th February 2005

III.1.4 UK Beam-System Meteor Radar

The United Kingdom meteor wind radar is a beam system and consists of two orthogonal, low-elevation antennas. The antennas were directed NW (300º) and NE

(30º) in the last 7 years of operation from April 1998 to December 2005 and the system was located near Castle Eaton (51.67º N, 1.78º W). The radar has previously been located at Sheffield (53.45º N, 1.58º W) from Jan 1988 to May 1995 and also at Rutherford Appleton Laboratory (51.57º N, 1.32º W) from March 1995 to April 1998. This gives a data set of almost 17 years.

The radar is a pulsed system which measures the winds using the Doppler echo technique similar to that of the SKiYMET, but only in the two beam directions within ± 15 degrees. It is assumed that the winds in the meteor region are moving in a bulk motion over the large area. These winds can then be used to calculate the average horizontal wind speeds in the zonal and meridional directions. A detailed description of the radar is given by Muller (1970) and Muller *et al*. (1995).

The radar operates at a transmission frequency of 36 MHz, with a peak pulse power output of 10 kW. Every other transmitted pulse is doubled as a rudimentary pulse coding to unambiguously determine the range. The two receiver antennas are six-element Yagi, which are located half a wavelength above the ground, to give a main beam radiation pattern of ~ 30° in elevation. Because the UK meteor radar does not have height finding equipment, echoes are assumed to be representative of the flow near the peak height of the meteor distribution. For this frequency (~ 30 MHz) this is ~ 90 km.

III.3 Satellites

Satellites that circle the globe are able to give us considerable insight into the Earth's atmosphere. They are often useful in studies of the lower and middle atmosphere. Lower atmosphere/weather satellites tend to have geostationary orbits, designed to maintain a position above a point on the equator. They are able to continuously study the regions below, but often have very poor views of the Polar Regions. Near Polar orbiting satellites, however, are able to observe a much larger area/volume and up to much higher latitudes. Here we will consider the Aura Satellite and the MLS instrument because it was chosen to be used in these studies.

III.3.1 Aura Satellite and the MLS Instrument

The Earth Observing System (EOS) Microwave Limb Sounder (MLS) is one of four instruments onboard the NASA Aura satellite (Schoeberl *et al.,* 2006), which was

launched on July 15th 2004. Aura is part of NASA's A-train constellation of Earth observing satellites (Figure III.12). These satellites travel in formation, with each satellite making measurements within a short time of each other. The Aura satellite is in a near-polar, 705 km orbit. It is also in a Sun synchronous orbit and therefore orbits with the same local time as the Earth rotates underneath.

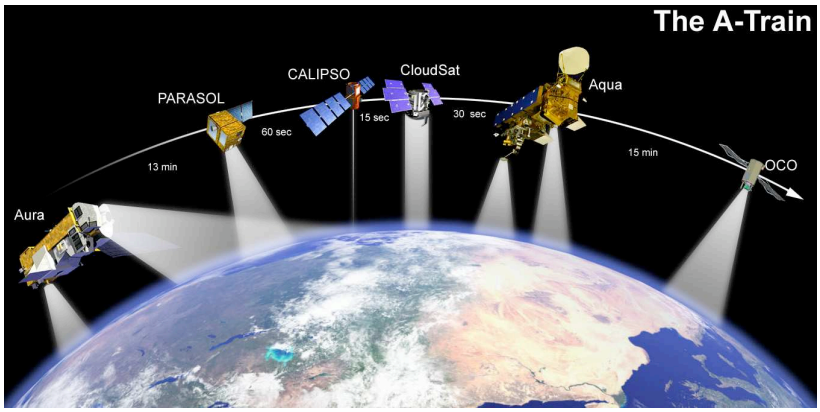


Figure III.12: An artist's concept showing how Aura flies in formation with the other Earth Observing Satellites (Image credit: NASA).

The MLS instrument observes naturally-occurring thermal microwave emission from many molecules in the Earth's limb. It views forward along the Aura spacecraft flight direction and scans its view from the ground to ~ 96 km every 25 seconds. Figure III.13 shows the Aura spacecraft and the MLS instrument. It is able to remotely sense vertical profiles of atmospheric gases, temperature, pressure, and cloud ice (Waters *et al.*, 2006).

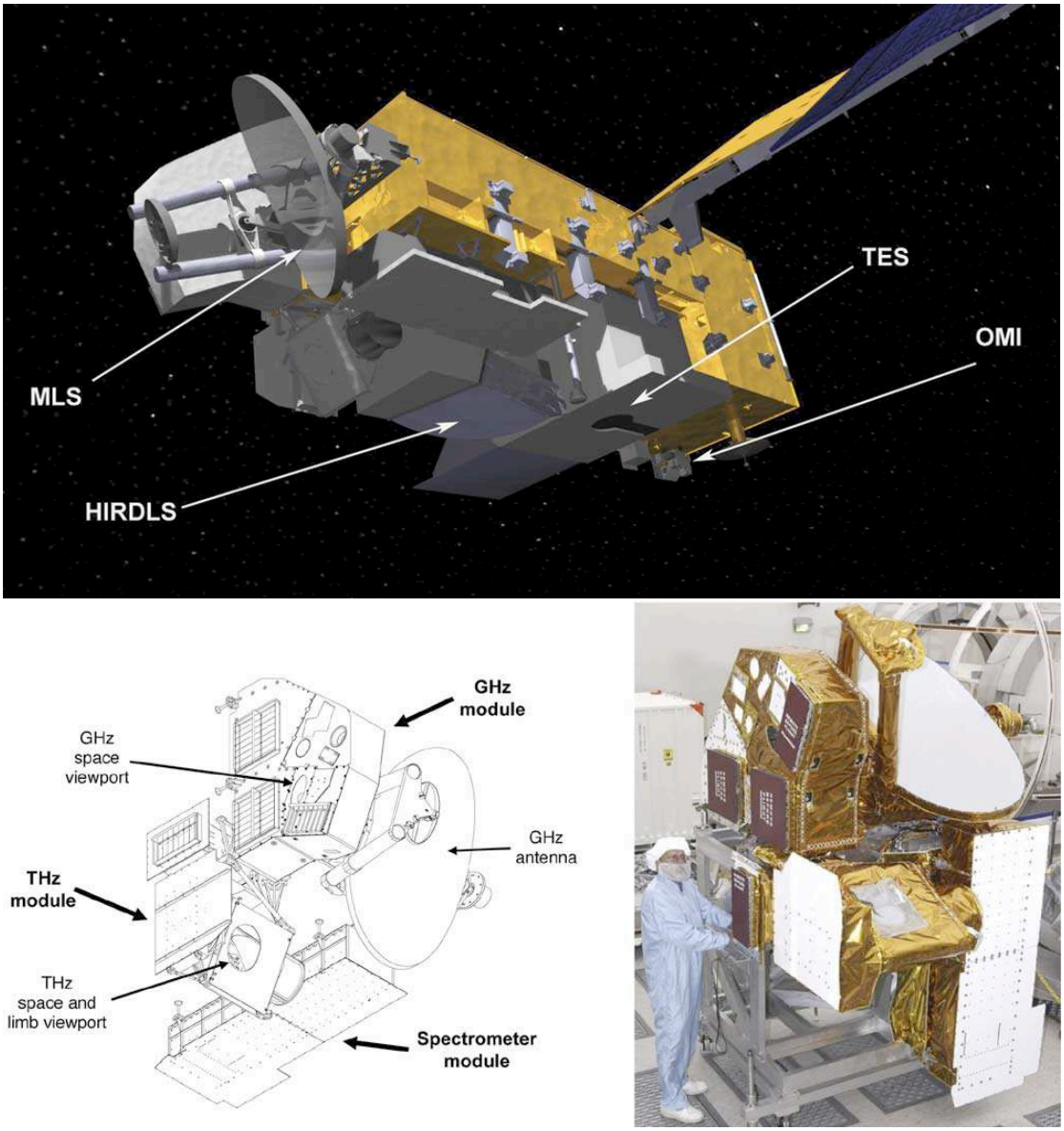


Figure III.13: Top: A model simulation of the Aura spacecraft. The locations of the four onboard instruments are indicated (After Schoeberl *et al.,* 2006). Bottom left: a line drawing of the MLS instrument (After Waters *et al.,* 2006). Bottom right: a photograph of the MLS instrument (After Waters *et al.,* 2006).

The first MLS instrument was aboard the NASA Upper Atmosphere Research Satellite (UARS) launched in Sept 1991. But due to power conservation and the lifetime of the instrument's systems, observations post March 1994 became intermittent and the last observations were made at the end of August 2001. The second MLS instrument is currently on the NASA Aura mission. The EOS MLS began making scientific observations on 13 August 2004.

The instrument measures ~ 3500 vertical profiles per day along the suborbital track. The vertical scan rate varies with altitude, with a slower scan providing greater integration time in the lower regions (0 - 25 km) as shown in Figure III.14. The MLS

vertical scans are synchronized to the Aura orbit such that vertical scans are made at essentially the same latitudes each orbit, with 240 scans performed per orbit. Therefore 240 Profiles of geopotential height are retrieved per orbit, spaced at 1.5° great circle angle along the suborbital track at an essentially fixed set of latitudes.

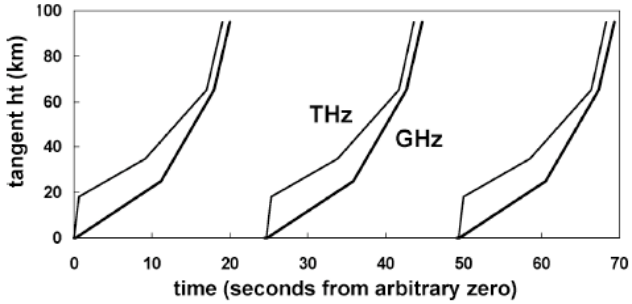


Figure III.14: An example of the MLS limb scanning profiles. High resolution scans are taken at heights below 20 km, with a lower resolution between 20 and 65 km and even lower over 65 km (After Waters *et al.,* 2006).

The instrument observes thermal microwave emission by the atmosphere in five spectral regions between 115 GHz and 2.5 THz. Geopotential height and temperature measurements are calculated using the radiances observations near the 118 GHz O_2 spectral line and the 234 GHz $O^{18}O$ spectral line.

The output contains temperature on 47 pressure levels and geopotential height on the 100 hPa reference level. Both the geopotential height and temperature are available as standard level 2 data products. Currently version 2.2 is the most recently available and will be used in this thesis. The geopotential height profiles are formed from the instrument data using a reference geopotential height at the 100 hPa level and hydrostatic integration of the temperature profiles (Schwartz *et al.,* 2008). The vertical resolution for the geopotential height is not well defined, as it is an integrated quantity, however, the calculated precision values of the MLS Level 2 Version 2.2 data are ~ 35m from 316 hPa to 100 hPa, ~ 45m at 1 hPa, ~ 110m at 0.001 hPa.

All of the standard Level 2 data products of the Aura satellite and their useable height ranges are shown in Figure III.15, the MLS's products are show in red. Each MLS Level 2 Version 2.2 data point is reported with a corresponding precision. A reference atmosphere is taken from the CIRA86 climatology (GEOS-5 below 1 hPa). If the retrieval does not improve precision by at least a factor of two on the *a priori* value, above 1 hPa, then it has failed to extract sufficient additional information from radiances. The retrieved values will then be significantly influenced by the *a priori* data. Such precisions are reset to negative. These data are not recommended for use in scientific studies and are therefore not used here.

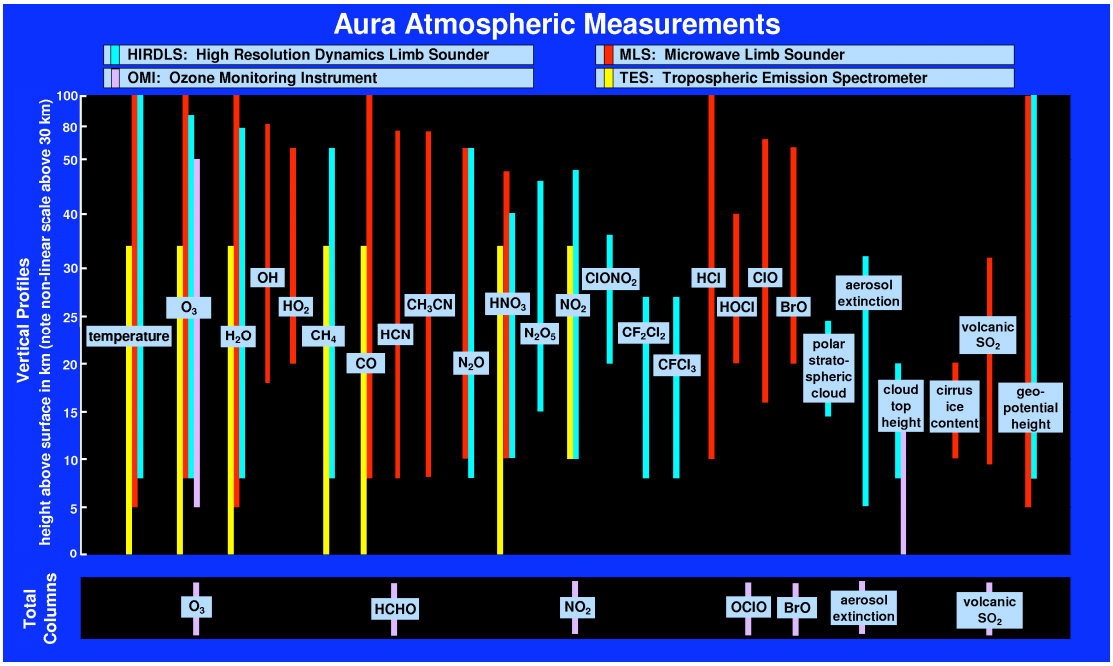


Figure III.15: The parameters measured by the instruments aboard the Aura satellite. MLS parameters marked in red. The lines indicate the useful range for individual profiles (After Schoeberl *et al.,* 2006).

Three additional data flags are provided for each vertical profile. The *status* field indicates operational abnormalities or problems with the retrievals. Profiles for which status is an odd number are not used as they are also not recommended for scientific study. Even values excluding zero, indicate that the profile has been marked as questionable by the data processing software. This is usually because of possible significant influence from thick clouds.

The *quality* field indicates the degree to which the measured radiances have been fit by the algorithms. Larger values of quality generally indicate better radiance fits, whereas values closer to zero indicate poor radiance fits, and less reliable data. Profiles having quality values less than 0.6 are unsuitable for scientific use and excluded. This removes ~ 4% of the temperature profiles. The *convergence* field is used to flag profiles which within a set of ~ 10 profiles failed to converge well. Convergence values around 1.0 typically indicate that the retrieval has converged. Profiles with convergences greater than 1.2, however, are unsuitable for scientific use and excluded. This typically rejects ~ 2% of the profiles or 0.5% in addition to those flagged in the quality field.

Temperature and geopotential height at pressures lower than 0.001 hPa (~ 97 km) or higher than 316 hPa (~ 9.4 km) are also not recommended for use in scientific studies and excluded.

III.6 Summary

This chapter introduced methods of remote sensing used in this thesis. These included the meteor radar technique and how it is used to make MLT region observations. Also described were the microwave limb-sounding instrument and its geopotential height data product. From the various techniques available to observe the middle atmosphere, these two were chosen to be used, as they are well suited for these studies.

The meteor wind radar is an ideal technique for measuring long term trends and large scale phenomenon of the MLT region. Other techniques such as lidar offer a higher time resolution; satellite observations offer better global coverage; imagers offer better spatial resolution, etc.. However, each of these techniques have their own drawbacks. Each technique offers different advantages and disadvantages over others. The choice of instrument to be used therefore depends on the type of question to be answered.

In conclusion, the meteor wind radar is a powerful tool for measuring and observing the dynamics of the MLT region. Combining observations with those of satellite techniques can prove invaluable in the study of large-scale phenomenon, such as mean winds, tides and planetary waves in the middle atmosphere.

Chapter IV

IV. Arctic Planetary Waves

Planetary waves are large-scale perturbation in the atmosphere that extend coherently around the full longitudinal circle. The most important planetary waves are the quasi-stationary mid-latitude Rossby normal modes. These are waves that correspond to the natural mode of variability of the Earth's atmosphere. They are the largest amplitude planetary waves and they propagate vertically upward from the troposphere. These waves are largest and most variable in the middle atmosphere during winter. Studying the behaviour of these waves is essential in understanding the coupling and dynamics in the atmosphere.

The waves were first identified in the atmosphere in 1939 by Carl-Gustaf Arvid Rossby who went on to explain their motion (Wallace and Hobbs, 2006). The waves in the mesosphere and lower thermosphere, which have received the most attention in the literature, are probably the "5-day", "10-day" and "16-day" waves, all with zonal wavenumber one and the "quasi-2-day" wave which typically has zonal wavenumber three. There are also a large number of studies containing equatorial travelling planetary waves (e.g., Kelvin waves).

These planetary waves can interact with other waves such as quasi-stationary Rossby waves, gravity waves, and tides; thereby affecting the dynamics of the middle atmosphere. The majority of published observations of the planetary wave field in the MLT region have been made by ground-based systems located in the Northern Hemisphere mid-latitudes. Here, planetary waves with periods between 1.5 and 20 days are considered.

IV.1 Esrange Data

The Esrange meteor radar was used to investigate planetary waves in the Arctic MLT region. Measurements to be analysed are of horizontal winds at mesopause heights recorded by the meteor radar at Esrange (68°N 21°E) as described in Chapter III. The Esrange meteor radar has operated continually from October 1999. Here, data from the interval October 1999 to November 2007 are considered and were analysed to yield hourly estimates of zonal and meridional winds in six height gates between 80 and 100 km.

Figure IV.1 presents a wavelet analysis from the hourly spaced horizontal winds of the Esrange meteor radar with periods between 1.5 and 20 days, for the year 2005. The wavelet transform has an advantage over the Fourier transform when representing functions which occur in bursts, have sharp peaks or discontinuities. The wavelet function breaks up a time series into different scale components and each component can then have a resolution to match its scale. The type of window function used was a Morlet wavelet, which is a sine wave consisting of six cycles at each scale, multiplied by a Gaussian envelope. This particular wavelet is ideal for analysis of data in which a wave occurs in bursts and has high amplitudes for only a few cycles of the wave. This type of activity is known to represents planetary wave activity very well. Everything above a 95% significance level is displayed, and the wavelet shows activity around: 10 and 16 days during the beginning and end of the year; 5-day wave activity is a little more frequent occurring in bursts throughout the year and 2-day wave activity in small bursts also throughout the year. The activity is not necessarily exactly on the specific period, and is more a range of values, for example the 10-day wave activity in 2005 is generally nearer to a period of ~ 8 days and also the quasi-2-day wave is named such as it is known to have a largely variable period.

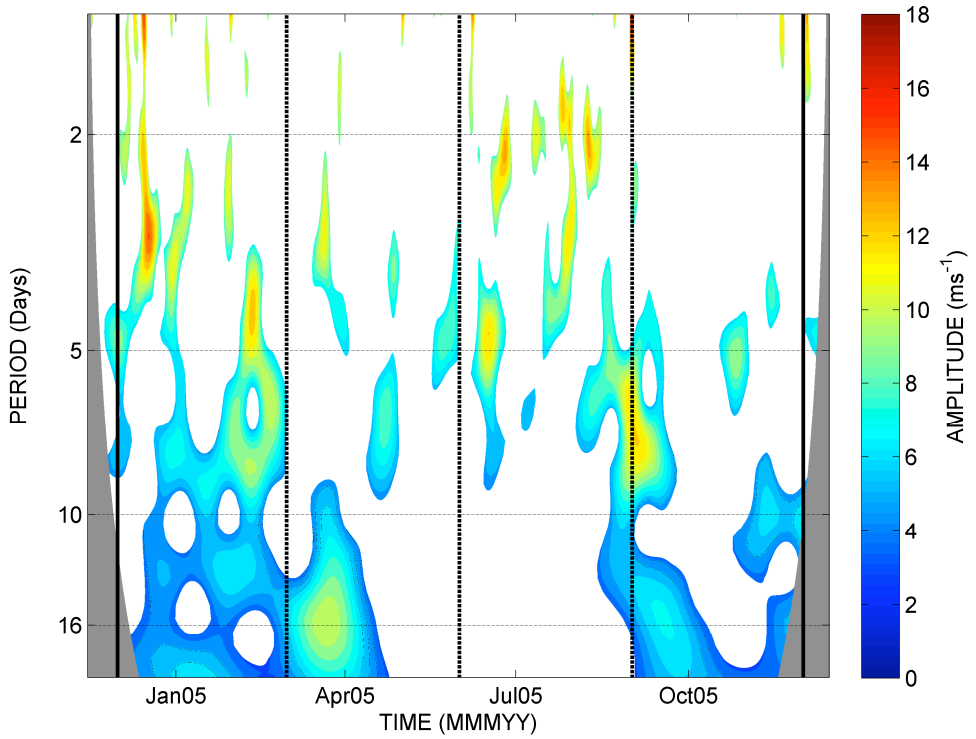


Figure IV.1: Planetary wave field of the meridional wind component over Esrange for the year 2005. Vertical solid black line indicates 1st December and the dashed lines indicate 1st March, 1st June and 1st September respectively to indicate season.

Because the activity of the planetary waves could be over a range of frequencies, the data was broken up into specific period bands as shown in table IV.1. These bands were chosen as they were centred on the predicted period of the corresponding planetary wave i.e., the predicted eigen-period for a specific planetary wave (Salby, 1984). To ensure that all of the energy from the planetary wave field was considered, the bands were contiguous, with no frequency/period gaps.

Table IV.1: The possible period range of the four most common planetary waves

Planetary Wave	Period Range (days)
"Quasi-2-day" Wave	1.5 – 3.4
"5-day" Wave	3.4 – 7.7
"10-day" Wave	7.7 – 13.0
"16-day" Wave	13.0 – 20.8

The horizontal wind data from the Esrange meteor radar was then band pass filtered using these period bands. The band pass used here was an elliptic filter. This filter type offers a steeper "roll-off" characteristic than Butterworth filters, however there is more ripple on the elliptical filter than a similar Butterworth filter, but this has however been kept to a minimum. In general, elliptic filters meet given performance specifications with the lowest order of any filter type. We assume that all of the

variance in this band is associated with the particular wave. The frequency response of the filter used to band pass the quasi-2-day wave is shown in Figure IV.2.

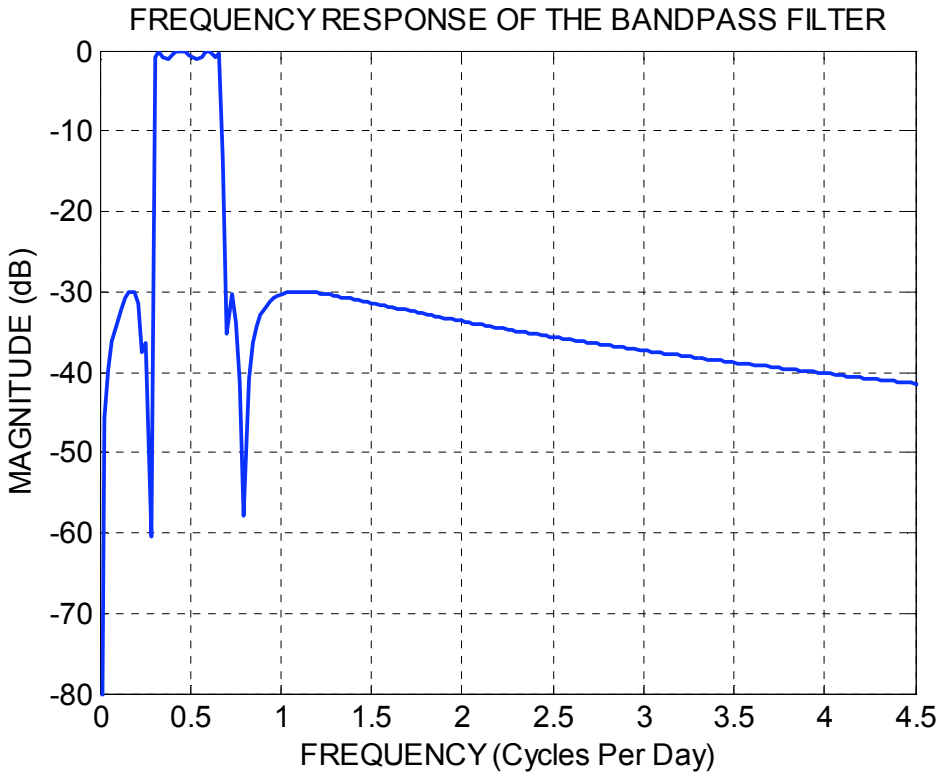


Figure IV.2: The frequency response to the elliptical filter used to band pass the quasi-2-day wave. The line indicates the amount attenuation at each frequency. This type of filter has high roll off characteristics indicated by the steep sides of the band, but more ripple over the band period than other similar filters.

Once the data had been band passed, a variance was then taken for each month of each of the bands. These variance values then provide a simple proxy of wave activity within this period band for each month. The variances for each month were then averaged over the seven years of data available from Esrange and the results of this analysis are presented in Figure IV.3a, b. This figure shows that the 2-day band and 5-day band generally have the largest variances and have a semi-annual cycle with maxima during the wintertime and summertime and minima at the equinoxes. For the 10-day band and 16-day band the majority of activity is found during the winter.

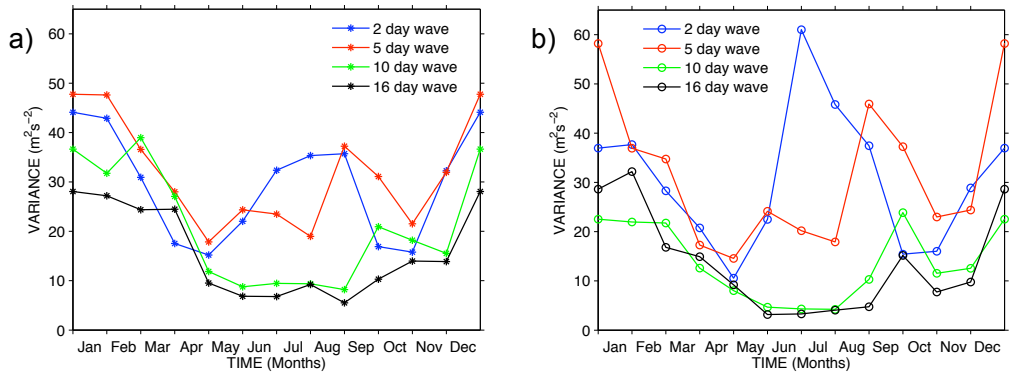


Figure IV.3: Composite year monthly variances for a) zonal and b) meridional Esrange meteor wind radar data over seven and a half years from 1999 - 2007. Data were first band-passed to identify the activity in the: (blue) quasi-2-day wave band; (red) 5-day wave band; (green) 10-day wave band and (black) 16-day wave band.

As a measure of the uncertainty in the band pass filtering and variancing of the data set, each of the hourly-calculated wind values was randomised in time. This gives a data set which has the same total variance as the original time series, but is unlikely to have any coherent wavelike oscillations. This data set is then band-passed as before for each of the planetary wave components. This in effect produces an estimate of the error introduced by manipulating the data when there are no genuine waves present. The composite variances can then be taken as a noise floor for the data and the results of this analysis are presented in Figure IV.4a, b. It should be noted that because the data has been randomised, this is likely to produce more high frequency components than lower ones. When the data is randomised from the coherent oscillations, with smooth transition from hour to hour it now becomes sharp random changes i.e., more high frequency components. This can be seen in Figure IV.4a and b as the 2-day wave variance is the highest, then the 5-day, the 10-day and then the 16-day is the lowest.

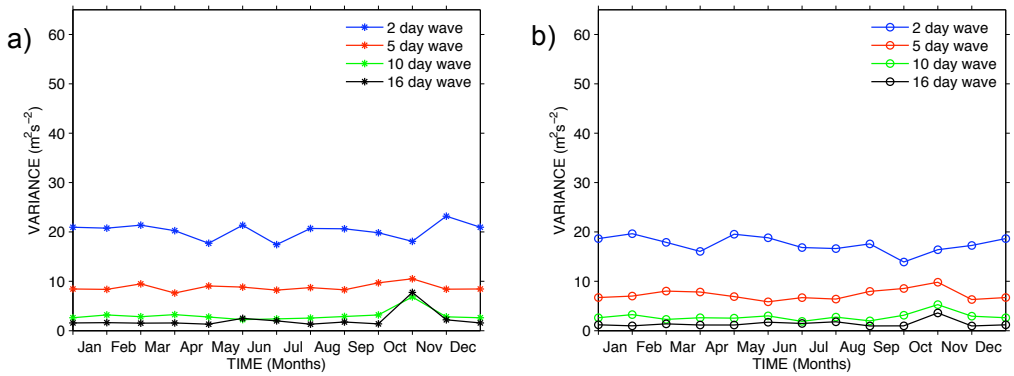


Figure IV.4: Composite year monthly variances for a) zonal and b) meridional **randomised** Esrange meteor wind radar data over seven and a half years from 1999 - 2007. Data were first band-passed to identify the activity in the: (blue) quasi-2-day wave band; (red) 5-day wave band; (green) 10-day wave band and (black) 16-day wave band.

In general, the noise levels of Figure IV.4a, b are smaller than the variances of Figure IV.3a, b. The lowest variances of Figure IV.3a and b are often comparable to their corresponding noise level of Figure IV.4a and b. This suggests that no significant wave amplitudes are present during this time; however, during times when the activity is higher than that of Figure IV.4a and b, the activity suggests that there are waves present.

IV.2 The Two-Day Wave

The quasi-two-day wave is one of the most conspicuous features of the mesosphere and lower thermosphere (MLT) region. At middle and low latitudes the quasi-2-day wave reaches maximum amplitudes of several tens of ms^{-1} in late summer, where it can be the largest-amplitude feature present, but it is largely absent at other times (e.g., Muller and Nelson, 1978). This quasi-2-day wave is usually only seen in the summertime. It is known to interact with and modulate the tides of the mesosphere (e.g., Teitelbaum and Vial, 1991 and Mitchell *et al*., 1996) and may also impact the photochemistry of this region (e.g., Kulikov, 2007). However, activity in the data from Esrange (68°N) presented in Figure IV.3 shows that in the quasi-2-day wave band, between 1.5 – 3.4 days, has a semi-annual cycle in wave activity.

The summertime quasi-2-day wave is believed to be primarily a westward propagating, zonal wavenumber 3 (i.e., three cycles of the wave around the latitude circle, hereafter W3 etc.) wave. This has been identified with the (3, 0) Rossby normal mode (Salby, 1981) and baroclinic instabilities, if there is sufficient eastward shear above the summertime westward jets (Plumb 1983). Significant W4 components and somewhat smaller W2 components have also been reported in both observations and in modelling studies (e.g., Meek *et al*., 1996; Norton and Thuburn, 1996; Pancheva *et al*., 2004).

In contrast to these mid-latitude observations, recent studies have revealed significant wave activity with a period near two days in the winter Arctic MLT region (Nozawa *et al*., 2003a, b; Manson *et al*., 2004). These studies used ground-based meteor and MF radars to reveal significant wave amplitudes in winter, but were unable to determine the zonal wavenumber or direction of propagation of the wintertime polar two-day wave because they had insufficient longitudinal coverage in the radars used. This wave will now be referred to as the wintertime polar two-day wave, so as not to confuse it with the quasi-2-day wave.

In the polar winter stratosphere, a two-day wave has been observed and identified as an eastward propagating, zonal wavenumber 2 (E2) planetary wave. This wave is believed to be one of a series of waves generated by instabilities in the polar night jet (e.g., Venne and Stanford, 1982; Hartmann, 1983; Prata, 1984; Lait and Stanford, 1988 and Manney *et al*., 1988). Nozawa *et al*. (2003b) suggested that the polar MLT-region wintertime two-day wave may be an E2 wave connected to these stratospheric oscillations – but this could not be confirmed in the absence of longitudinal measurements.

Palo *et al*. (2007) reported an E2 wave in the mid-latitude, summer, southern-hemisphere MLT region occurring simultaneously with a W3 wave. They attributed the E2 wave to a non-linear coupling between the W3 quasi-2-day wave and the migrating diurnal tide.

Here, we investigate the nature of the wintertime polar two-day wave observed in the MLT region by an Arctic meteor radar and in data from the MLS instrument aboard the NASA Aura satellite. In particular, the longitudinal coverage provided by the satellite observations allows us to determine zonal wavenumbers and the extended height coverage will also help to reveal the relation of this MLT-region wave to two-day waves in the stratosphere. Finally, we address the mechanisms whereby this wave might be excited (e.g., non-linear coupling or excitation by barotropic instability).

IV.3 Data Analysis

Data from two independent sources are used in this study. The first source is measurements of horizontal winds at mesopause heights recorded by the meteor radar at Esrange (68°N 21°E). This radar measures the winds with typical height and time resolutions of about 3 km and 1 hour, respectively (See Chapter III, Mitchell *et al*. (2002) for details of the Esrange meteor radar and Hocking *et al*. (2001) for a description of the technique). Here, again, data from the interval October 1999 to November 2007 are considered and were analysed to yield hourly estimates of zonal and meridional winds in the six height gates between 80 and 100 km.

The second source of data is the microwave limb sounding (MLS) instrument aboard the NASA Aura satellite - part of the A-train constellation of satellites (described in Chapter III and by Waters *et al.,* 2006). Data from Aura over the interval August 2004 to July 2008 are considered here. MLS yields geopotential height as one of its level-2, version 2.2 data products covering the height range of ~ 10 – 97 km. The Aura satellite

orbits at 705 km in a Sun-synchronous orbit and so passes through only two local times at any given latitude. The latitudinal limb measurements of the satellite with respect to local time are shown in Figure IV.5. At 68°N, the latitude of the Esrange meteor radar, these local times are about 3.1 and 12.3 hours. The highest latitude accessible to the satellite is just over 80°. Aura's limb sounding instruments are all designed to observe roughly along the orbit plane and the MLS is on the front of the spacecraft in the forward velocity direction.

The *geopotential* $\Phi(z)$ at any point in the Earth's atmosphere is defined as the work that must be done against the Earth's gravitational field to raise a mass of 1 kg from sea level to that point (z). Therefore, the geopotential $\Phi(0)$ at sea level $(z=0)$ is by definition taken as zero. *Geopotential height* is defined by the equation:

$$Z = \frac{\Phi(z)}{g_0} \qquad \text{IV.1}$$

Where g_0 is the globally averaged acceleration due to gravity at the Earth's surface (Wallace and Hobbs, 2006). Therefore taking the geopotential height at a certain pressure level can be considered as variations of that level and used to measure the activity of the atmosphere. From this, perturbations due to waves can be observed in a time series of geopotential height.

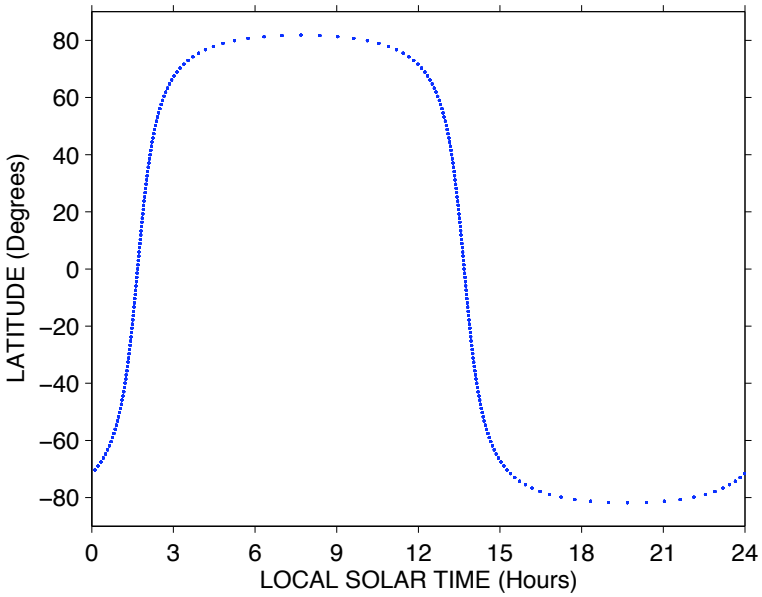


Figure IV.5: EOS (Earth Observing System) Aura Satellite orbital track in local time, with an equatorial crossing time of 1:45 ± 15 minutes.

Figure IV.6 shows a similar variation in the satellite's orbital track as in Figure IV.5, however, this figure shows how the track varies with longitude in one day. The track covers a large number of longitudes and has the best longitudinal coverage at high latitudes. The satellite takes ~ 16 days to complete its entire path before it begins covering the same grid again.

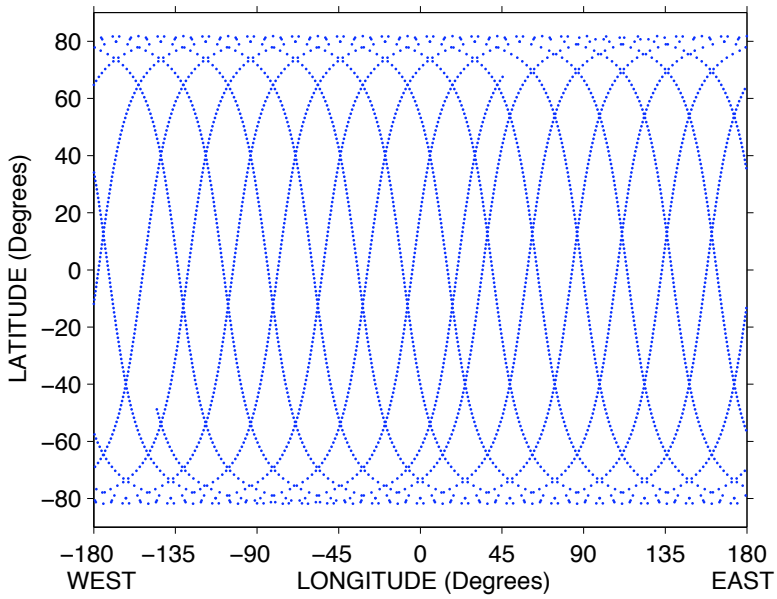


Figure IV.6: EOS (Earth Observing System) Aura Satellite orbital tracks observed on the 1st January 2005.

Here, we have applied the data analysis method of Wu *et al*. (1995a), to the geopotential height data. In this method, the geopotential height data is divided into latitude bands of 10 degrees width (e.g., 60°–70°N). A data set is then assembled consisting of geopotential height values for successive passes of the satellite through each band. Because of the path of the satellite, the universal times and longitudes of the crossings through any given latitude band are not independent and, in fact, form a regular matrix. An example of this matrix is shown in Figure IV.7 for data taken between 63° and 73°N.

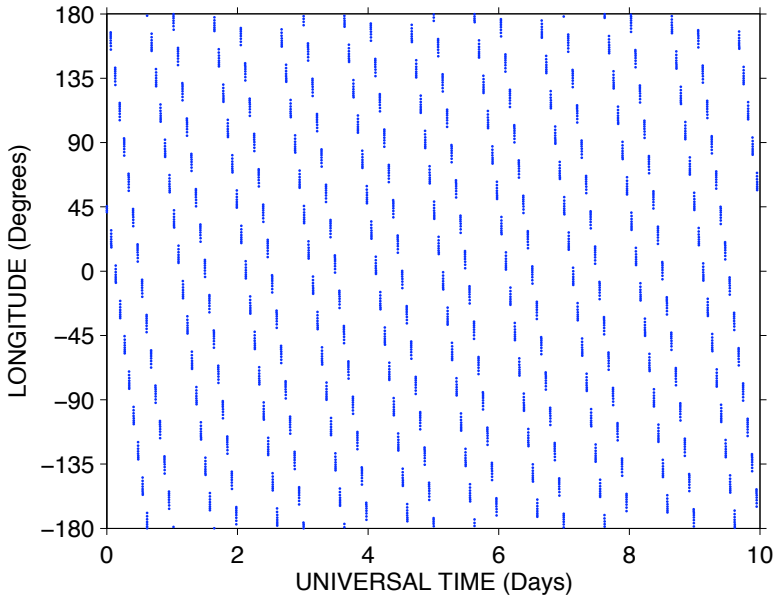


Figure IV.7: Data coverage for satellite pass between 63° – 73°N from 1st – 11th January 2005 forming a regular matrix between time and longitude.

As the matrix is very regular, this can lead to aliasing. Therefore, to bring this to a minimum, only 75% of the data is actually used and this data is taken at random points, as randomness in this matrix can help to reduce or even eliminate the aliasing (Wu *et al*., 1995a, b).

Planetary waves are assumed to vary sinusoidally in time and longitude. An example of this is shown in Figure IV.8a, b. These show the longitude against time variation of a) an eastward propagating zonal wavenumber 2 two-day wave with a period of 48 hours and b) a westward propagating zonal wavenumber 3 two-day wave also with a period of 48 hours. By taking a vertical slice of Figure IV.8a, this wave has two maxima and two minima in the full longitudinal circle i.e., wavenumber 2; Figure IV.8b has three maxima and three minima i.e., a wavenumber 3. Figure IV.8a (b) travels east (west) with time and has a maximum and minimum every two days, seen in a horizontal slice.

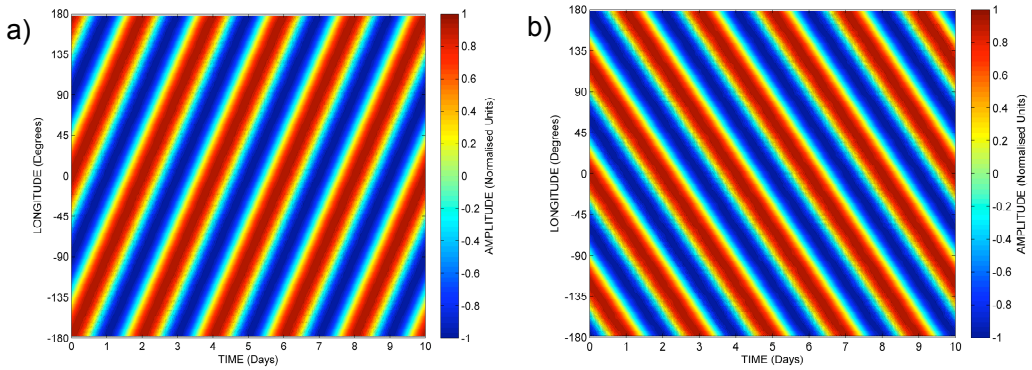


Figure IV.8: Model normalised variation in longitude and time for a) an eastward zonal wavenumber 2 two-day wave and b) a westward propagating zonal wavenumber 3 two day wave.

The sinusoidal functions of zonal wavenumbers from 4 to -4 (i.e., 4W to 4E) and periods of 42 to 56 hours were then least-squares fitted to the two dimensional matrix. This yields geopotential amplitudes and phases for waves with any combination of these wavenumbers and periods within a particular latitude band. Note that positive wavenumbers correspond to westward propagating waves and negative to eastward propagating waves. Also note that because there are only two local times within any latitude band, this analysis is unable to resolve tides without complications caused by severe aliasing. The least-squares fit is a particularly useful technique when using unevenly sampled data or data with many gaps and an ideal choice for fitting sine waves to the randomly sampled two-dimensional satellite data.

The spectral resolution of this data is given using the method of Wu et al. (1995), who give the equation:

$$\left(\Delta\sigma T\right)^2 + \left(\Delta s\right)^2 = a^2 \qquad \text{IV.2}$$

Where $\Delta\sigma$ is the frequency resolution, Δs the wavenumber resolution and T is the total sample length. Wu et al. (1995) gives the value of the constant a as normally being ~1, however, they recommend values between 1 and 1.45. Using a = 1.2, as Baumgaertner *et al.* (2008) did in a similar analysis and using the sample length of 15 days as in the analysis presented here, the spectral resolution is ~ 6.6 hours.

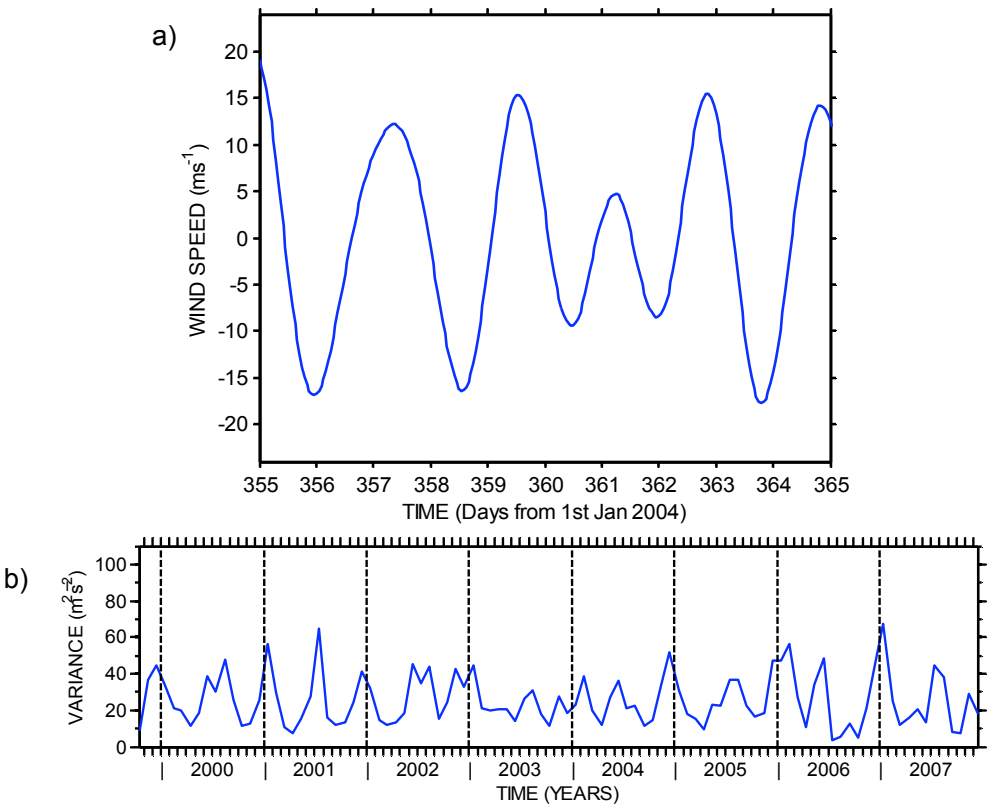
The radar data were analysed to produce a climatology of waves with a period near two days. The wide filter used in Figure IV.2 was reduced to include only known quasi-2-day wave periods of between 1.5 and 2.6 days. For a particular height gate, the horizontal winds were band-pass filtered to reveal oscillations with periods between the 1.5 and 2.6 days, using an elliptical filter. The transfer points of the filter used are at 1.49 and 2.61 days for the 99% and 1.41 and 2.76 days for the 1%. We assume that all of the variance in this band is associated with the two-day wave. For each month of

data a variance value was calculated for each height gate. These variance values then provide a simple proxy of wave activity within this period band.

IV.4 Results

As an example of the wintertime 2-day wave, Figure IV.9a presents band-passed zonal winds over Esrange recorded at heights of 89 to 92 km during 10 days in December 2004. In the figure it can be seen that there are approximately five complete cycles of the wave occurring in a 10 day segment; therefore during this time the wave has a period of about 48 hours.

Figure IV.9b and c presents monthly variance values for zonal and meridional winds, respectively, measured over Esrange from a height gate spanning 89 to 92 km from October 1999 to November 2007. The figure clearly shows that variance tends to maximise in winter and summer and has equinoctial minima, i.e., there is a semi-annual cycle. There is also considerable inter-annual variability. For instance, the summertime maximum varies from values as small as about 30 m^2s^{-2} in 2003 (zonal component) and 2004 to values as large as about 100 m^2s^{-2} in 2002 (meridional component).



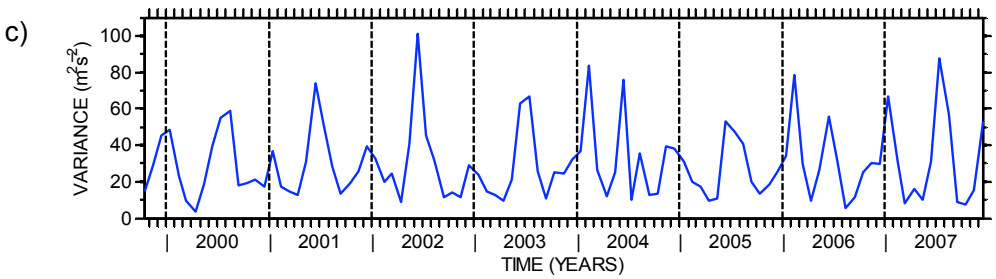


Figure IV.9a-c. Band-passed horizontal winds from the Esrange meteor radar from heights of 89 to 92 km. The data have been band passed between periods of 1.5 and 2.6 days. a) Zonal component for 10 days in December 2004 showing five cycles of a wave in 10 days. b) and c) Monthly variances of the band-passed horizontal winds for b) zonal winds and c) meridional winds. The dashed vertical lines indicate the boundary between each year.

This semi-annual behaviour is in marked contrast to that observed at middle and low latitudes where wave activity with a period near two days maximises in the summer and is significantly weaker in the winter (e.g., Muller *et al*., 1978).

To reveal a representative seasonal cycle, a composite-year analysis was carried out using the data of Figure IV.9. Figure IV.10 presents this composite year analysis for the zonal and meridional components. The semi-annual cycle of variance is very apparent. In the case of the winter maximum, the zonal and meridional variances are about the same size and are ~ 40 m^2s^{-2}. However, in the case of the summer maximum the meridional variance is significantly larger than the zonal, some ~ 68 m^2s^{-2} c.f. some ~ 35 m^2s^{-2}. In both zonal and meridional cases the equinoctial minimum have values of about ~ 15 m^2s^{-2}.

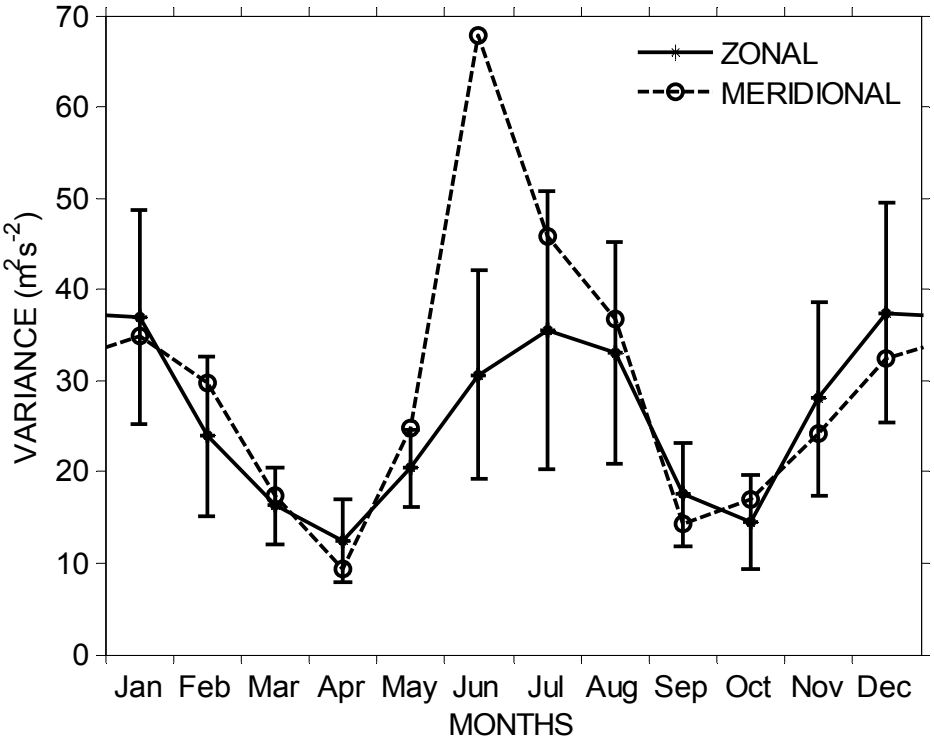
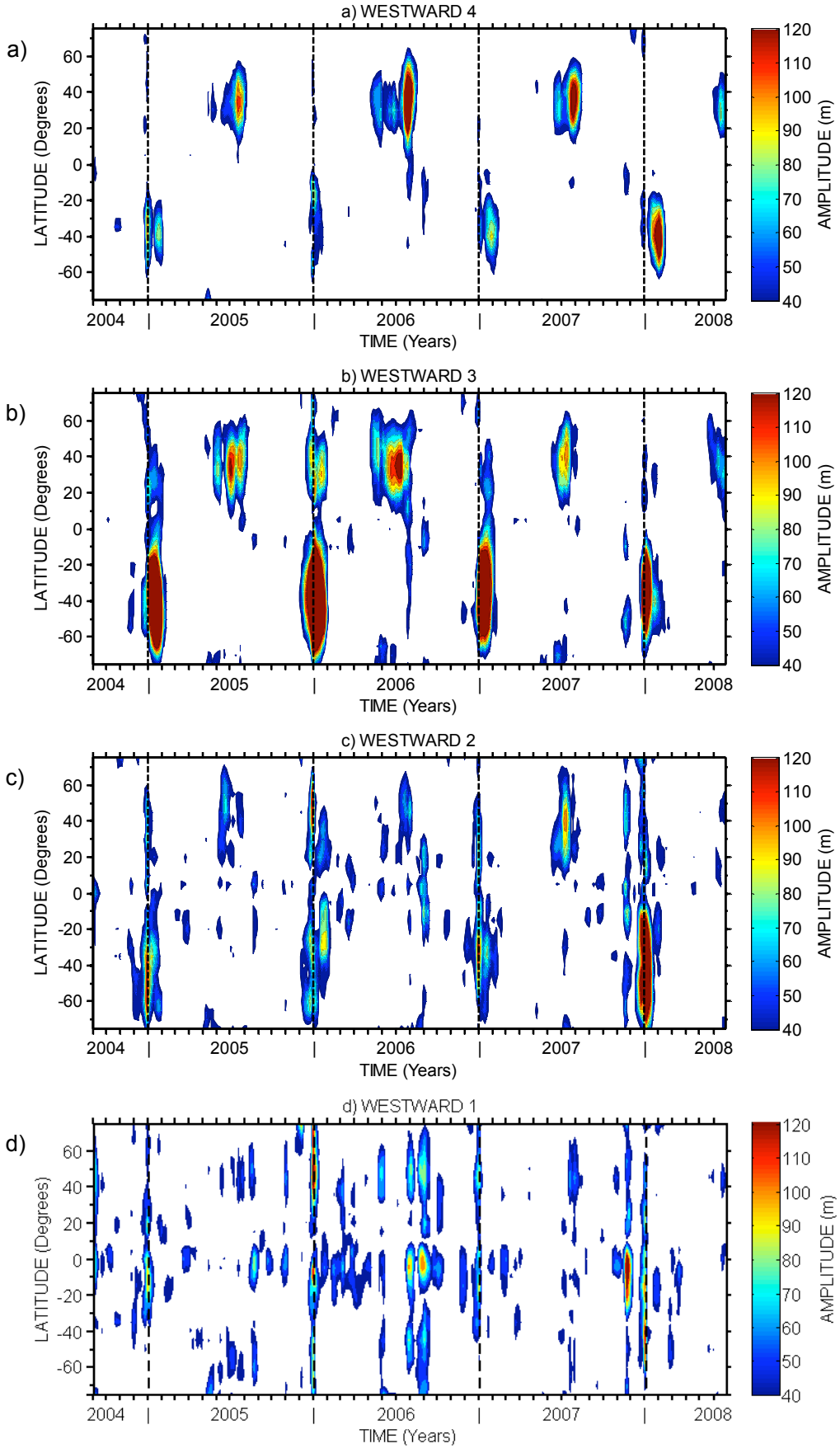


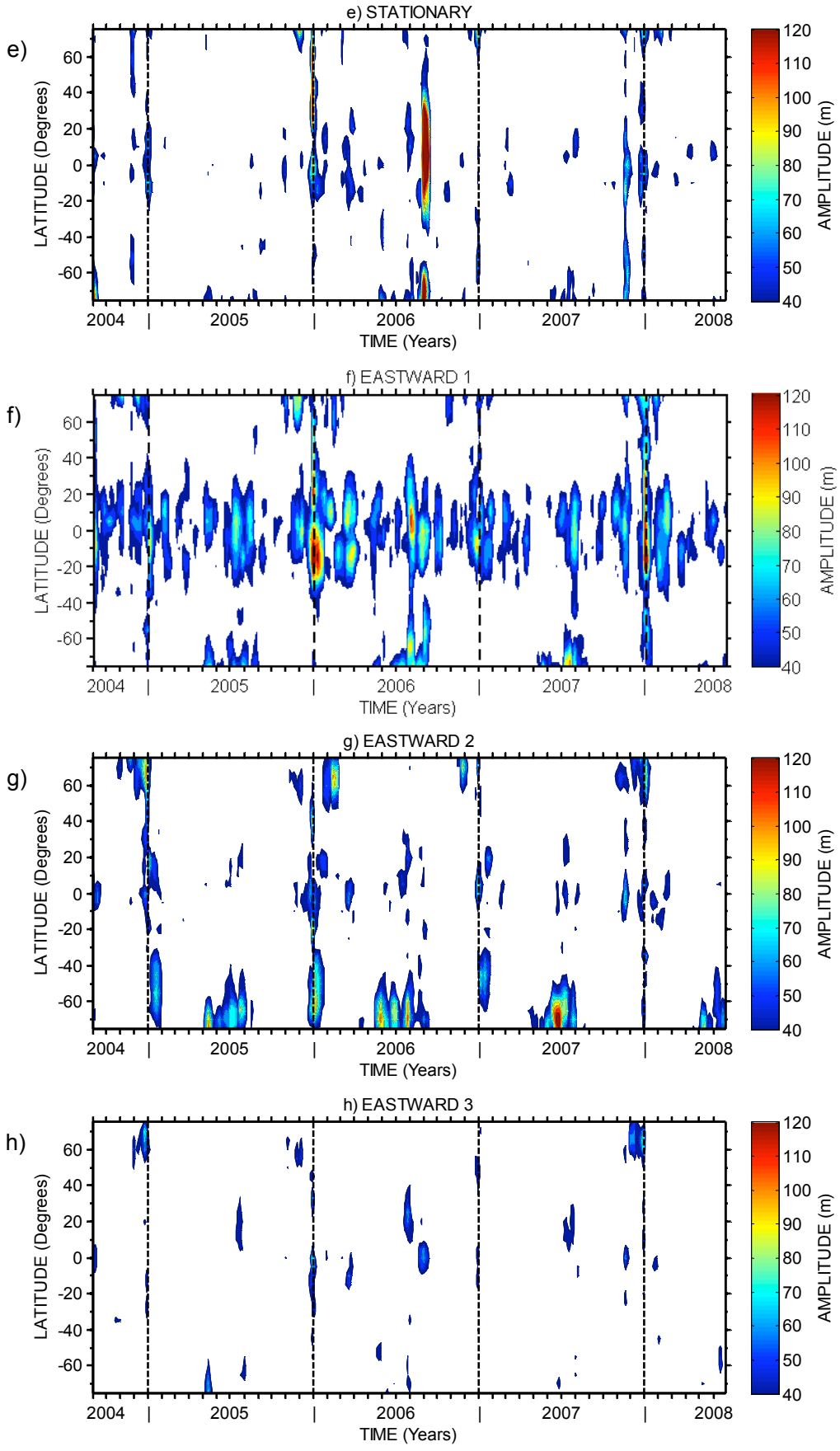
Figure IV.10: A composite-year analysis of the data of Figure IV.9ab and c. The solid line indicates results for the zonal winds and the dashed line meridional winds. Error bars indicate the standard deviation and are marked only on the zonal component.

These results suggest that there is considerable wave activity with a period near two days present in the Arctic winter and summer. This reinforces the conclusions of Nozawa *et al*. (2003a, 2003b) and Manson *et al*. (2004) who also note wintertime activity at a period of two days at high northern latitudes.

The MLS data provide an opportunity to investigate the zonal wavenumber(s) and global context of these wave features evident in the summer and winter Arctic mesosphere. For each 10 degree latitude band from 80°N to 80°S, the average geopotential height amplitude was calculated over the period range 42 – 56 hours. These averages were calculated using a time window of 15 days incremented through the data set in steps of 5 days. The calculation was performed for zonal wavenumbers W4 to E4.

The result of this analysis is an estimate of wave amplitude against latitude and time for each of the wavenumbers W4 to E4 over the interval August 2004 to November 2007. Figure IV.11 a-i presents these results for zonal wavenumbers a) W4, b) W3, c) W2, d) W1; e) The stationary; f) E1, g) E2, h) E3 and i) E4 at a height of about 90 km (selected to match the radar observations). In each Figure, the contours are plotted only for amplitudes larger than 40 m.


a) WESTWARD 4
a)
LATITUDE (Degrees)
60
40
20
0
-20
-40
-60
2004
2005
2006
2007
2008
TIME (Years)
AMPLITUDE (m)
120
110
100
90
80
70
60
50
40
b) WESTWARD 3
b)
LATITUDE (Degrees)
TIME (Years)
AMPLITUDE (m)
c) WESTWARD 2
c)
LATITUDE (Degrees)
TIME (Years)
AMPLITUDE (m)
d) WESTWARD 1
d)
LATITUDE (Degrees)
TIME (Years)
AMPLITUDE (m)


e)
e) STATIONARY
LATITUDE (Degrees)
60
40
20
0
-20
-40
-60
2004
2005
2006
2007
2008
TIME (Years)
AMPLITUDE (m)
120
110
100
90
80
70
60
50
40
f)
f) EASTWARD 1
LATITUDE (Degrees)
TIME (Years)
AMPLITUDE (m)
g)
g) EASTWARD 2
LATITUDE (Degrees)
TIME (Years)
AMPLITUDE (m)
h)
h) EASTWARD 3
LATITUDE (Degrees)
TIME (Years)
AMPLITUDE (m)

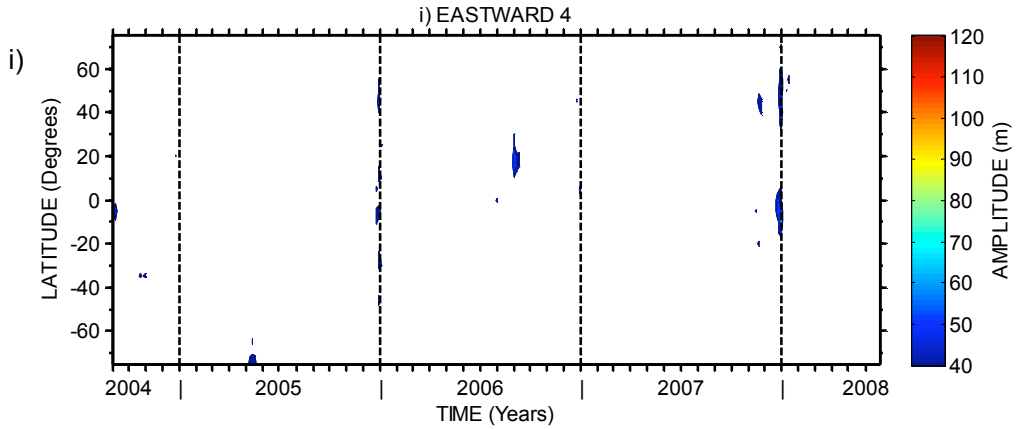


Figure IV.11 a-i: Geopotential height amplitudes as a function of time and latitude for Aura MLS data at a height of ~ 90 km, for oscillations with periods between 42 and 56 hours and different zonal wavenumbers. a) Westward 4, b) westward 3, c) westward 2 d) westward 1, e) stationary, f) eastward 1, g) eastward 2, h) eastward 3 and i) eastward 4. Amplitudes below 40 m are not plotted and dark red areas indicate amplitudes greater than 120m. The solid vertical black lines indicate the boundary between years.

We will now consider each figure in turn. The W4 results of Figure IV.11a reveal a very clear structure which repeats from year to year. The amplitudes maximise in each hemisphere at mid-latitudes during the summer. Wave amplitudes are strongest during the months of June to August in the northern hemisphere and December to February in the southern hemisphere. Amplitudes tend to reached values of ~ 140 m in the summer, except in summer 2006 in the northern hemisphere which reached > 200 m. Note that there are no significant amplitudes (> 40 m) in winter in either hemisphere. Further, note that the amplitudes are small at latitudes poleward of ± 60° in each hemisphere.

The W3 results of Figure IV.11b also reveal a very clear structure with amplitudes which also maximise in each hemisphere at mid-latitudes during the summer. In this case, however, the summertime southern-hemisphere amplitudes are significantly larger than in the northern hemisphere. Wave amplitudes are strongest during the months of June to August in the northern hemisphere and January and February in the southern hemisphere. In the northern hemisphere, amplitudes reach values of ~ 140 m. However, in the southern hemisphere, peak amplitudes exceed 250 m in each of the three summers observed. Note again that there are no amplitudes > 40 m in winter in the southern hemisphere. There are, however, bursts of activity in the northern hemisphere at mid-latitudes during each winter. These bursts reach amplitudes as large as 90 m, but are in all cases smaller than the summertime amplitudes. Again, note that the amplitudes are small at latitudes northwards of 60°. However, in the southern hemisphere, significant amplitudes are present to latitudes as far south as –75°.

The W2 results of Figure IV.11c reveal a rather similar picture to that of the W3 and W4, however, less well defined and lower amplitudes. Again, although there are amplitude maxima in each hemisphere during the summer, there is also activity of equivalent amplitude in the northern hemisphere winter. In the southern hemisphere, wave amplitudes are strongest during the months of December to February. The amplitudes reach values of ~ 100 m in both the northern and southern hemisphere. The oscillations extend as far polewards as ± 70° during the summer.

Results for wavenumbers W1, 0, E3 and E4, Figures IV.11d, e, h, and i respectively, are seen to have rather smaller amplitudes generally < 50 m and the results tend to be quite scattered. The amplitudes of E3 and E4 yielded negligible amplitudes which are generally less than 40 m.

The E1 and E2 results of Figure IV.11f and g reveal a rather different picture. In contrast to the other wavenumbers, the amplitude of E1 and E2 maximises in both hemispheres during the winter. For the E2, there is also significant activity in the southern hemisphere summer, which may be an effect of aliasing from the W3, since these maxima occur at the same time as those of the W3 (see Section IV.4.1). The E2 activity generally occurs poleward of ± 50° and the E1 poleward of ± 60°, however the E1 also has a lot of activity near the equator between ± 30°. The E2 wave amplitudes are strongest during the months of November to February for the northern hemisphere and June to August plus January for the southern hemisphere. Similar is also seen in the E1. Amplitudes reach values of ~ 100 m in both hemispheres for the E2 and ~ 80 m in both hemispheres for the E1.

To see if these observations can explain the behaviour observed over Esrange (Figure IV.9), the data were reanalysed to examine the latitudes from 63° to 73°N at a height of ~ 91 km. The monthly variances of two-day wave activity as observed by the radar are shown in Figure IV.12a for August 2004 to February 2007, the interval for which Aura MLS data are considered. Figure IV.12b, c presents the corresponding time series of geopotential heights over this latitude range. The E2 and E1 results are shown in Figure IV.12b and the W4, W3 and W2 results in Figure IV.12c. Results for wavenumbers W1, 0, E3 and E4 are not shown because these yielded negligible amplitudes.

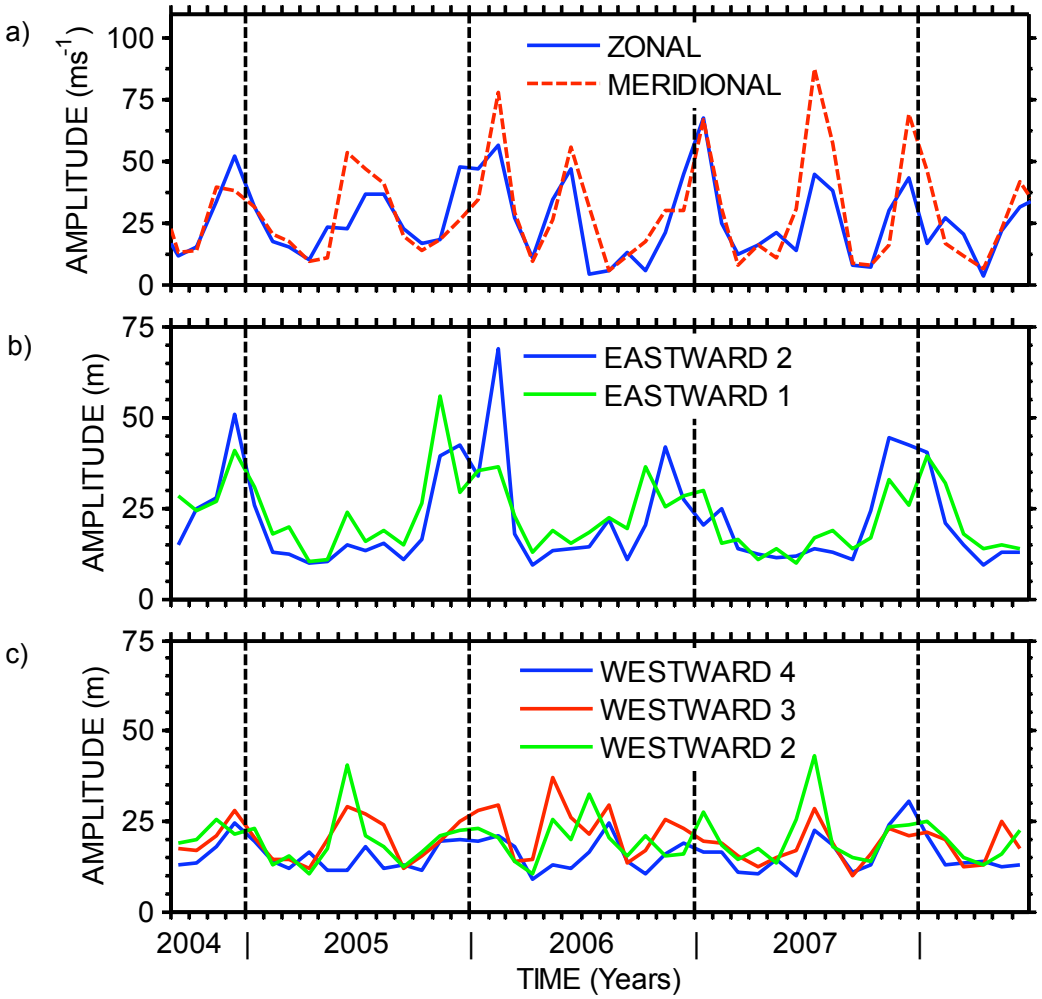
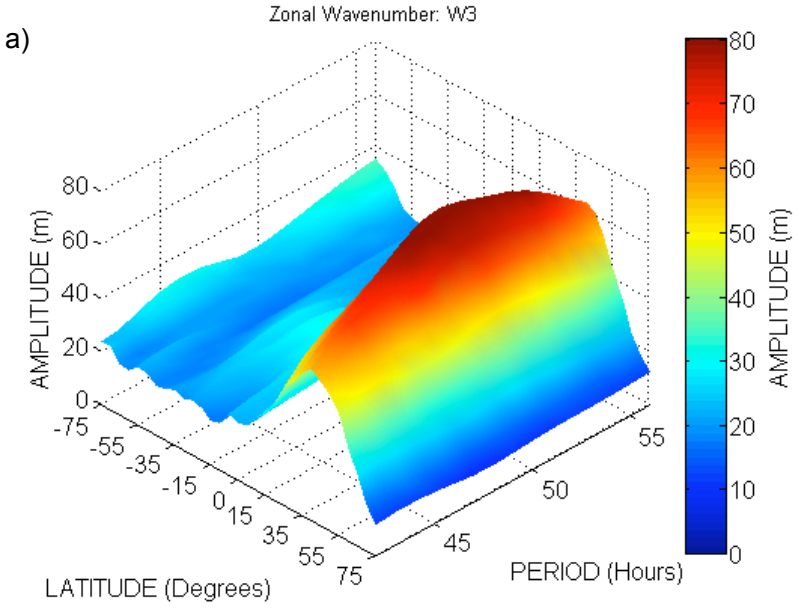


Figure IV.12: a) Horizontal wind variances from the Esrange meteor radar for a height of 89-92 km as per Figure IV.9. b) Geopotential height amplitudes at a height of ~ 91 km, between 63º and 73ºN, over the period range 42 to 56 hours for E1 and E2. c) as b) but for W2, W3 and W4.

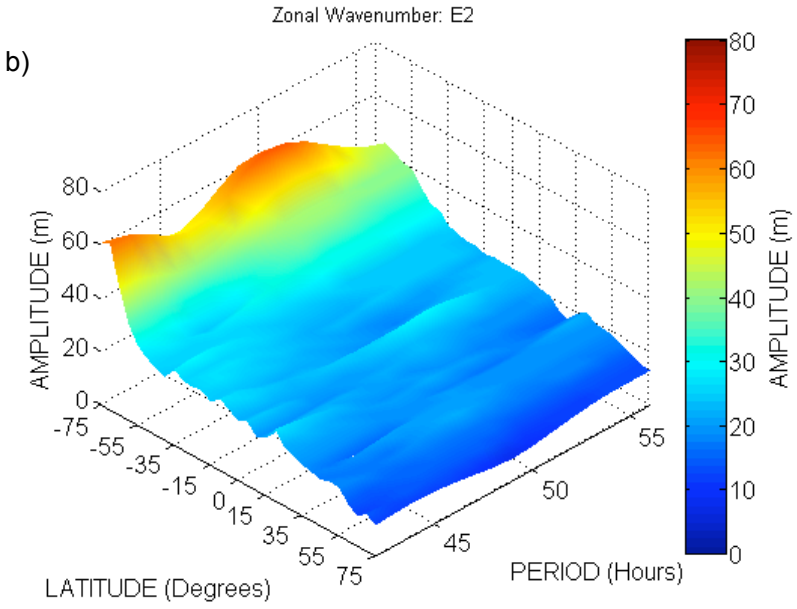
During winter, the greatest activity (geopotential amplitude) is present in the E2 component, closely followed by the E1; there is some activity in the W3. However, because of the effect of aliasing we do not believe this wintertime W3 activity is significant (see Section IV.4.1). The summertime activity seems to be a combination of the W3 and W2, with only small contributions from the W4, E2 and E1. These results show that the summer time and winter time polar two-day wave have very different wavenumbers, i.e., largely W3 and E2, respectively. Figure IV.13a, b and c, show the structure of the W3, E2 and E1 waves respectively in period and over latitude. The W3 has the majority of its activity in the summer hemisphere at middle latitudes and is generally quiet at other times. The maximum wave activity generally occurs around ~ 50 hours. The E2 however tends to increase in amplitude towards the winter pole and is likely to maximise at higher latitudes than those shown here. The period of this wave seems to have a maximum near ~ 50 hours and another of shorter period maximising at < 42 hours. The E1 is similar to that of the E2 with the peak amplitudes likely being

further poleward than 80° and a period closer to 3 days. It is also possible to see that there is a large amount of E1 activity near the equator, possibly connected with ultra-fast Kelvin waves, as it has the classic exponential decrease away from the equator, from the change in the Coriolis force. Note that the poor wave period resolution is due to the data being taken in 15 day means. Taking a longer period would have the effect of smearing out the burst of activity and 15 days was found to be the optimum.

a)



b)



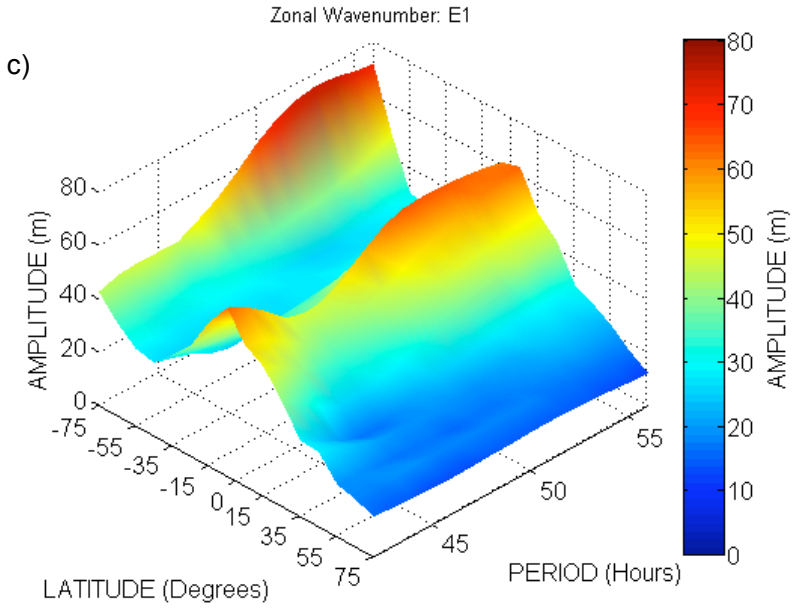


Figure IV.13: The structure of the a) W3, b) E2 c) E1 waves at ~ 90 km during northern hemisphere summer (JJA) for August 2004 to February 2007, between 80°S and 80°N and for the periods 42 to 56 hours based on 15 day means.

Prata (1984) showed that there was a wintertime, E2, polar two-day wave in the stratosphere. To determine if this wave is connected to that observed here in the mesosphere, we examined the MLS results from 10 to 97 km and 63° to 73°N. Figure IV.14 presents time-height contours of the geopotential amplitude of the E2 two-day wave calculated as 15-day means. The figure clearly shows that the wave observed in the mesosphere is also present to heights below 30 km and reaches maximum amplitudes at heights of about 60 km, i.e., near the stratopause. This means that the MLT-region and stratospheric observations are almost certainly observations of a single E2 planetary wave that is present at heights from below 30 km to above 100 km.

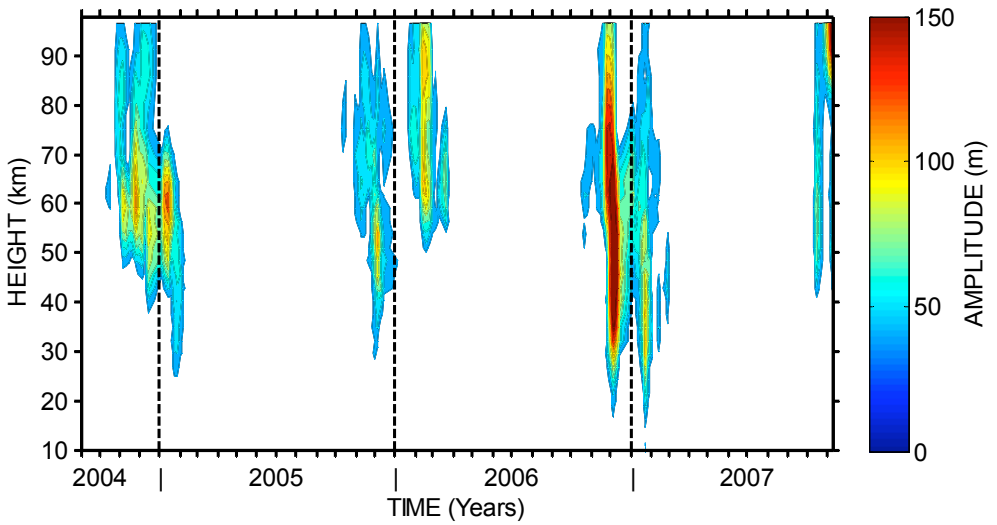


Figure IV.14: Time-height contours of geopotential-height amplitudes for an E2 wave of period 42-56 hours. The latitude range is 63° and 73° N.

The wintertime occurrence of this E2 wave can be explained in terms of the Charney and Drazin theorem (1961). The zonal phase speed, c_x, of the E2 two-day wave at the 68°N latitude of Esrange is ~ 52 ms^{-1}. i.e., the distance travelled divided by the time taken for the wave to propagate. The distance travelled is equal to:

$$dist = 2\pi\left(radius\right) \qquad \text{IV.3}$$

Where,

$$radius = R_e \cos\left(\varphi\right) + h \qquad \text{IV.4}$$

Where, R_e is the radius of Earth, h is the height measured above the surface and φ is the latitude. The radius at 68°N is much less than the radius of the Earth as shown in Figure IV.15.

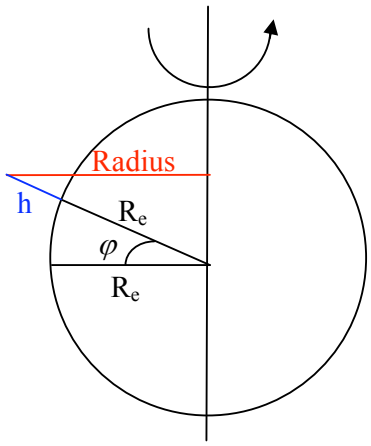


Figure IV.15: Schematic showing calculation of the radius at a height, h above the Earth's surface at a latitude of φ.

The time taken for the wave to propagate is the wavenumber multiplied by the period of the wave. The phase speed, c_x, of the wave is then the distance travelled in one circle of the Earth divided by the time taken for the wave to propagate round the full longitudinal circle.

Charney-Drazin theorem states that from the vertical structure equation of planetary wave motion:

$$k_z^{\,2} = \frac{\omega_B^2}{f^2}\left[\frac{\beta}{\overline{u} - c_x} - \left(k_x^{\,2} + k_y^{\,2}\right)\right] - \frac{1}{4H^2} \qquad \text{IV.5}$$

For vertical wavenumber, $k_z^{\,2} > 0$ the equation has internal wave solutions (i.e., vertically propagating), while for $k_z^{\,2} \le 0$ the solutions are external waves (i.e., vertically trapped). Therefore for stationary waves (i.e., a wave with zero phase speed ($c_x = 0$)) vertical propagation is possible only when the mean zonal winds satisfy (as in Section II.4.3):

$$0 < \overline{u} < U_c \qquad \text{IV.6}$$

Where,

$$U_c \equiv \frac{\beta}{k_x^{\,2} + k_y^{\,2} + \dfrac{f^2}{\omega_B^2}\left(\dfrac{1}{4H^2}\right)} \qquad \text{IV.7}$$

is the Rossby critical velocity. The criterion can also be extended for travelling planetary waves if eastward wind ($\overline{u} > 0$) is replaced by eastward wind with respect to the wave $\left(\left(\overline{u} - c_x\right) > 0\right)$ to give:

$$0 < \overline{u} - c_x < U_c \qquad \text{IV.8}$$

Therefore travelling planetary waves can only propagate for values of c_x such that $0 < \overline{u} - c_x < U_c$, where $\overline{u}$ is the zonal mean wind and U_c is the critical speed (Charney and Drazin, 1961; Holton, 1980 and Forbes, 1995).

For the latitude of Esrange (68°N) and assuming an E2 wave of meridional wavenumber 2, the critical speed, U_c is ~ 28.6 ms^{-1}. Therefore the E2 two-day wave should only be able to propagate in regions where the zonal wind speed lies between ~ 52 and 80.6 ms^{-1}.

To see if this is actually the case in our observations, we use the zonal mean zonal winds from the United Kingdom Met Office (UKMO) stratospheric assimilated data (Swinbank and O'Neill, 1994). The stratospheric assimilated dataset consists of a global analysis of three-dimensional winds, temperatures and geopotential heights. This data is routinely built, with a grid size of 2.5° latitude by 3.75° longitude, over 22 pressure levels from the ground to ~ 55 km. Observations of the middle atmosphere are assimilated into a global atmospheric general circulation model. Operational meteorological observations from many different sources are used, such as: satellite; radiosonde and aircraft measurements. The time period of data used is for the October 2004 to February 2007 interval corresponding to the Aura MLS dataset. It also covers the height range 10 to 65 km before March 2006 and 10 to 55 km after March 2006.

Also zonal mean winds from the Esrange meteor radar between 80 and 97 km, have been taken, apart from a month gap in August 2006. To match these to the MLS observations, they were calculated as 15-day means shifted by 5 days and within the 63° – 73°N latitude range for the UKMO stratospheric assimilated data. Figure IV.16 shows the zonal mean zonal winds from this analysis and also zonal mean winds from the Esrange meteor radar in the upper part of the figure. Note that the two datasets are not strictly comparable as the UKMO stratospheric assimilated winds are zonal means and the Esrange winds represent a single location.

From Figure IV.16 and IV.17, it can be seen that there is a strong correspondence between the strength of the zonal winds and the occurrence of the E2 wave. In particular the wave only reaches large amplitudes when the zonal winds are above about 50 ms^{-1}. To investigate this further, Figure IV.17 presents the amplitude data of Figure IV.14, with regions marked where the UKMO stratospheric assimilated winds have values in the range 52 to 81 ms^{-1}. It can clearly be seen that the E2 wave has largest amplitudes at heights and times that correspond to zonal winds falling within this range. This relation to the mean wind provides compelling support for the interpretation of the E2 two-day wave as a planetary wave. Note that the shaded regions of Figure IV.16 do not extend above ~ 65 km because this is the upper limit of the UKMO stratospheric assimilated analyses (~ 55 km between March 2006 and May 2007) and none of the Esrange mean winds meet the criteria (equation IV.6).

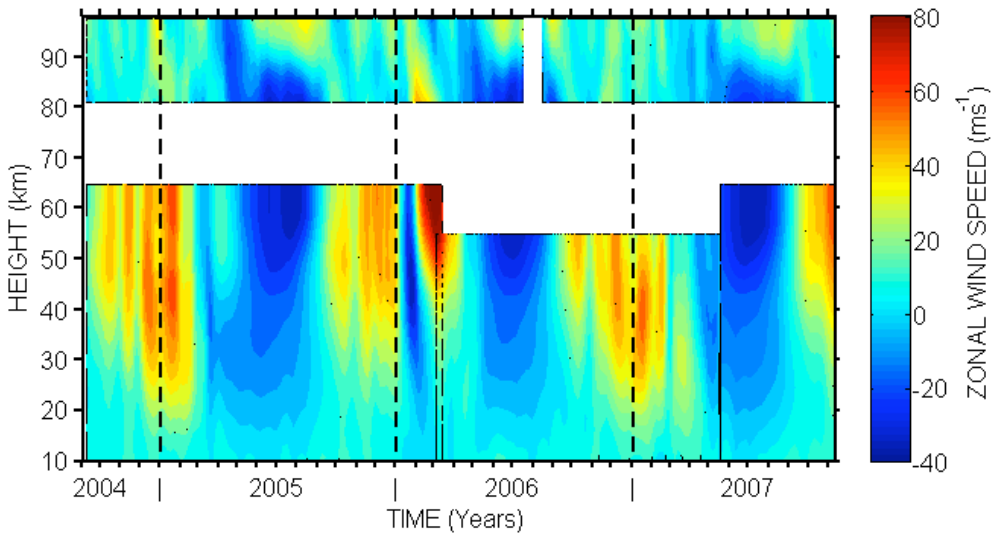


Figure IV.16: Time-height contours of the zonal mean zonal winds from the UKMO Stratospheric Assimilated Data bottom sections, and the mean zonal winds from the Esrange meteor radar in the upper sections.

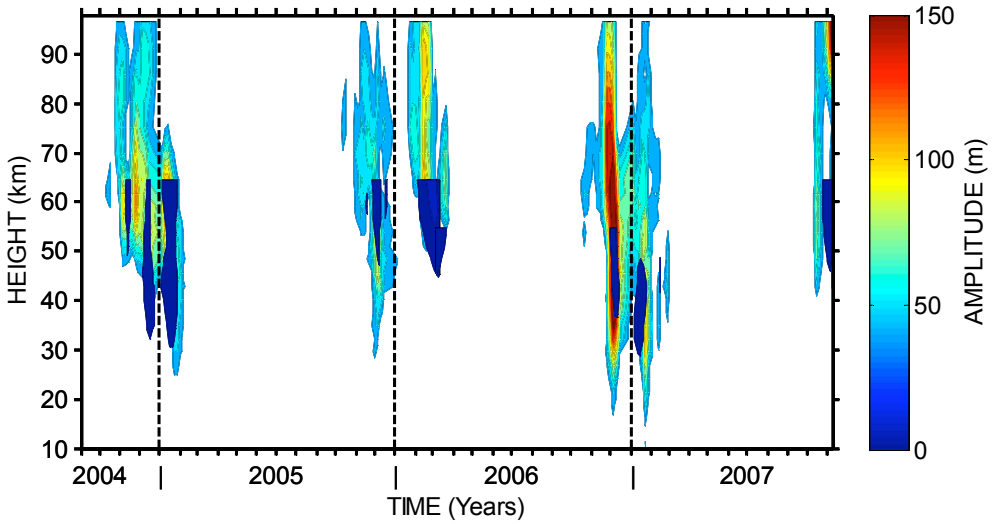


Figure IV.17: Time-height contours of geopotential-height amplitude same as Figure IV.14. The shaded regions indicate where the zonal mean zonal winds from Figure IV.16 satisfy the Charney-Drazin theorem for the E2 two-day wave. These regions predicted where this wave would be able to propagate.

Further, the occurrence of major sudden stratospheric warmings has the effect of reversing the zonal mean winds at high latitudes. Therefore, when a major stratospheric warming is present the E2 two-day wave should not be able to propagate as the wind regime will not satisfy the Charney-Drazin criterion. A major warming involves a reversal of the winds from the pole to 60° North in the northern hemisphere (at the 10 hPa level). Figure IV.18 shows latitude-time contours of the UKMO stratospheric assimilated zonal mean zonal winds at about 54 km, where the wave would otherwise have large amplitude. Two wind reversals associated with warmings are apparent in Figure IV.18 during March 2005 and January 2006 (indicated on the

figure). At the corresponding times in Figure IV.14, the E2 two-day wave is either not present or has small amplitudes. (However, see Section IV.5 for comments on the effect of stratospheric warmings on the excitation of the wave).

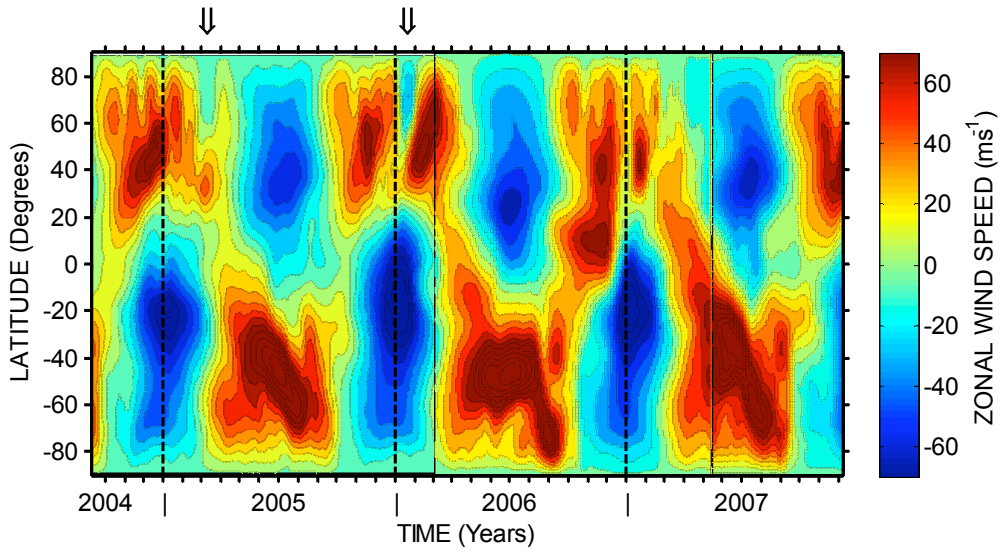


Figure IV.18: Time-latitude contours of zonal-mean zonal winds from the UKMO stratospheric assimilated data at a height of about 54 km. Arrows at the top of the plot indicate times of major stratospheric warmings.

Finally, the vertical gradients of phase in the MLS observations were used to calculate the vertical wavelength of the polar wintertime E2 two-day wave. Typical vertical wavelengths were found to be 123 ± 9 km.

From the above analyses, we conclude that the polar wintertime mesospheric two-day wave is dominated by an E2 two-day wave planetary wave, which extends through the stratosphere and mesosphere.

IV.4.1 Aliasing

Satellite studies of planetary waves can be complicated by aliasing because the longitude and time of observation are not independent (e.g., Wu *et al.*, 1995a). The relationship is:

$$\lambda_i = \lambda_a - ct_i, \qquad i = 1, 2, ..., N \qquad \text{IV.9}$$

Where, λ_a is the initial longitude of the ascending crossing points, c is the speed at which the crossing points move westward, this is equal to 1 day^{-1} for a Sun-synchronous satellite (i.e., Aura). The least squares fitting assumes that a particular

wave varies sinusoidally with $\sigma t_i + s\lambda_i$. Therefore, using equation, IV.9, the wave also varies as:

$$\sigma t_i + s\lambda_i = (\sigma - cs)t_i \qquad \text{IV.10}$$

So, as long as $(\sigma - cs)$ remains the same, the least-squares fitting will not be able to distinguish the different waves. Therefore, signals with periods and wavenumbers (σ_1, s_1) and (σ_2, s_2) are aliased to each other if:

$$\sigma_1 - cs_1 = \sigma_2 - cs_2 \qquad \text{IV.11}$$

The data can, therefore, have the same pattern present for multiple wavenumber-period combinations. In particular, a W3 two-day wave could alias to an E2 two-day wave and *vice versa*. For the Aura satellite, $c = 1$ day^{-1} as it is Sun-synchronous, therefore, a W3 with period and wavenumber (0.5, 3) would give aliased signals at the wavenumber – period combinations in Table IV.2. Note that the waves with negative period are travelling backwards in time and would therefore appear to travel in the opposite direction, i.e., the wave with period -0.5 cycles per day and wavenumber 2 would appear in the data as an eastward 2 (E2), rather than westward with period 0.5 cycles per day.

Table IV.2: Wavenumber – period combinations for an aliasing of a W3 two day wave.

Wavenumber, s_2	Period, σ_2 (Cycles per day)
6	3.5
5	2.5
4	1.5
2	-0.5
1	-1.5

The aliasing is further subcategorised according to the distance in the spectrum to the aliased wave. The waves (1.5, 4) and (-0.5, 2) are the closest and would show the maximum amount of aliasing. These waves would show up as a W4 with period 16 hours and an E2 two-day wave. The satellite data is not used for periods less than one day, therefore only the appearance of an E2 two-day wave could cause a problem. The aliasing of these waves could disappear if both the ascending and descending nodes are used in the calculation and there is a 12 hour local time difference between the ascending longitude and descending longitude i.e., 180°. For the Aura satellite, this is only true at the equator. At the latitude of Esrange (68°N) the ascending and descending nodes cross at 03:06 and 12:18, giving a difference of 9.2 hours.

The magnitude of the signal produced by aliasing at the closest frequencies depends on the difference in longitude between the ascending and descending nodes of the satellite's orbit. The magnitude can be calculated analytically in the regular sampling case and is proportional to:

$$\left|\cos\left(\pi\left(\lambda_a - \lambda_d\right)\right)\right| \qquad \text{IV.12}$$

It can be shown using the method described by Wu *et al*. (1995a) that in our observations for the latitude of Esrange (68°N), up to 35.84 % of the amplitude of the E2 two-day wave could be aliased into the W3 two-day wave, and vice versa.

An example of this aliasing is shown in Figure IV.19, which displays the measurement a satellite would take in its descending node with varying time, latitude and longitude for three cases in which there is only one wavenumber present, e.g., westward 3, eastward 2 or westward 1. These waves have been modelled as in Figure IV.8. The figure clearly shows that the measurements of the westward 3 and eastward 2 are very similar, with just a phase offset, whereas the westward 1 has a much larger period. Therefore, the least squares fitting analysis picks up a "ghost" signals at the aliased frequency when only one is truly present. This is a drawback of the technique. It can be reduced by randomising the data, but it still has an effect. Other examples of this are given in Figures IV.20b and IV.20b.

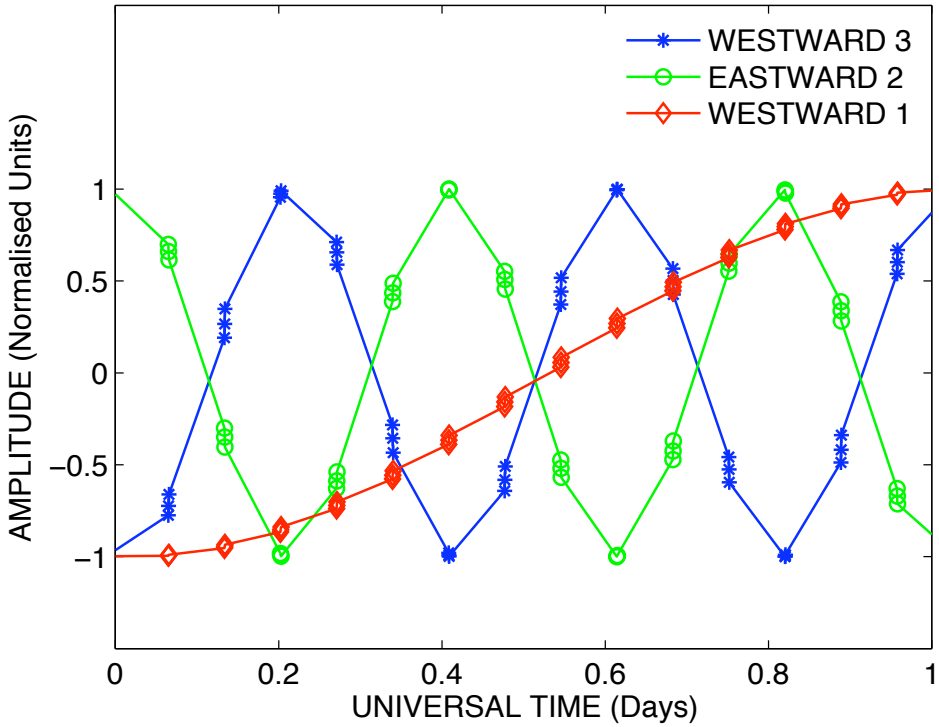


Figure IV.19: Three cases of the measurements a satellite would take during its descending node with varying time, latitude and longitude.

IV.5 Arctic Planetary Waves Discussion

In the polar winter stratosphere, a two-day wave has been observed and identified as an eastward propagating, zonal wavenumber 1 and 2 (E1 and E2) planetary wave. The E2 wave is believed to be one of a series of waves generated by instabilities in the polar night jet (e.g., Venne and Stanford, 1982; Hartmann, 1983; Prata, 1984; Lait and Stanford, 1988 and Manney *et al*., 1988). Prata (1984) reported observations of an E2 two-day wave in the polar winter stratosphere. Prata's observations were only made at a height of ~ 45 km. Our radar and satellite observations reveal that this wave likely extends to at least 100 km and that maximum amplitude occurs at ~ 60 km.

Nozawa *et al*. (2003b) also studied waves with periods near two day in the polar mesosphere from two radar locations at Tromsø (70°N, 19°E) and Poker Flat (65°N, 148°W). It was found that the activity was stronger in winter than in summer over both locations and, from the limited longitudinal coverage, it was noted that the times of maximum between the two sites were out-of-phase suggesting an even wavenumber. i.e., zonal wavenumber 2 or 4 were much more likely than the classical zonal wavenumber 3 of the Rossby normal mode. Nozawa *et al*. also suggested that the wintertime two-day wave could be an E2 wave connected to the stratospheric oscillations observed by Prata (1984), however, with the absence of longitudinal measurement coverage and an extended height range of observations, this could not be confirmed.

Palo *et al*. (2007) reported an E2 two day wave in the mid-latitude, summer, southern-hemisphere MLT region occurring simultaneously with a W3 wave. Palo *et al*. attributed the E2 wave as a non-linear coupling between the W3 quasi-2-day wave and the migrating (westward 1) diurnal tide.

In this mechanism a non-linear coupling produces secondary waves, including waves which have frequencies and wavenumbers that are the sum and difference of those of the primary wave and tide. This mechanism can therefore generate as a "difference" wave a wave that has E2 wavenumber and two-day period. However, the E2 polar two-day wave observed here is unlikely to have been generated by the non-linear mixing of a W3 and a diurnal tide because the W3 amplitudes that we observe can be explained entirely as a result of aliasing. The particular case that Palo *et al*. observed is that of Figure IV.20a. Figure IV.20a are observations taken at the same time as those of Palo *et al*. (i.e., 22nd – 29th January 2005) using geopotential height from the Aura satellite. Palo *et al*. used temperature measurements from the SABER (Sounding of the Atmosphere using Broadband Radiometry) instrument aboard the TIMED

(Thermosphere, Ionosphere, Mesosphere, Energetics and Dynamics) satellite, which give very similar results.

Palo *et al*. argued that non-linear coupling between two parent waves, the diurnal tide and the quasi-2-day wave, would produce a child wave which corresponded to the E2 two-day wave.

i.e.,: Parent Waves:

Diurnal Tide: $\left(f_{p1}, s_{p1}\right) = \left(1, -1\right)$

Quasi-2-day wave: $\left(f_{p2}, s_{p2}\right) = \left(0.47, -3\right)$

Child Waves:

First: $\left(f_{c1}, s_{c1}\right) = \left(f_{p1} + f_{p2}\right), \left(s_{p1} + s_{p2}\right) \rightarrow \left(f_{c1}, s_{c1}\right) = \left(1.47, -4\right)$

Second: $\left(f_{c2}, s_{c2}\right) = \left(f_{p1} - f_{p2}\right), \left(s_{p1} - s_{p2}\right) \rightarrow \left(f_{c2}, s_{c2}\right) = \left(0.53, +2\right)$

Where f is the frequency of the wave, s is the wavenumber, p indicates parent wave and c the child wave.

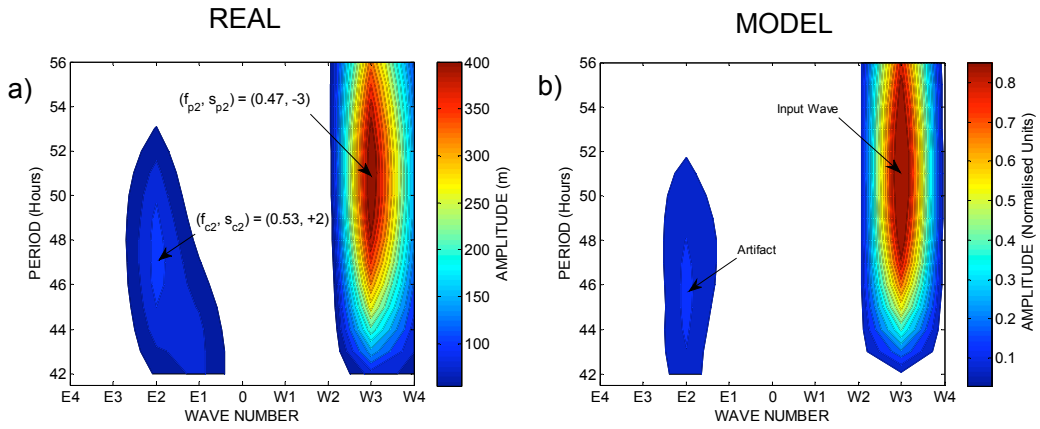


Figure IV.20a: Geopotential height observations of a W3 quasi-2-day wave occurring at the same time as an E2 two-day wave from 22nd – 29th January 2005 at a height of about 80 km and a latitude of 45°S. b) Model results from the least-squares fitting with only a W3 wave present in the model input data.

The second Figure, IV.20b shows the results of the same two dimensional least-squares fitting, but for model data rather than the real geopotential height data. The model data only included a W3 wave (i.e., the $\left(f, s\right) = \left(0.47, -3\right)$). This figure shows that aliasing of the W3 wave to the E2 period is possible and gives very similar results to that of the effect seen by Palo *et al.,* in Figure IV.20a.

A second example of this is given in Figure IV.21a and b. Figure IV.21a shows the geopotential height results from the 1st – 20th December 2004 at a height of about 90 km and a latitude of 68°N. This figure shows two E2 waves which could be non-linearly coupling with the migrating diurnal tide to produce the two observed W3 waves. Via the mechanism:

Green Parent Waves:

Diurnal Tide: $(f_{p1}, s_{p1}) = (1, -1)$

Two-day wave: $(f_{p2}, s_{p2}) = (0.55, +2)$

Green Child Waves:

First (Not Shown): $(f_{c1}, s_{c1}) = (1.55, 1)$

Second: $(f_{c2}, s_{c2}) = (0.45, -3)$

Red Parent Waves:

Diurnal Tide: $(f_{p1}, s_{p1}) = (1, -1)$

Two-day wave: $(f_{p2}, s_{p2}) = (0.46, +2)$

Red Child Waves:

First (Not Shown): $(f_{c1}, s_{c1}) = (1.46, 1)$

Second: $(f_{c2}, s_{c2}) = (0.54, -3)$

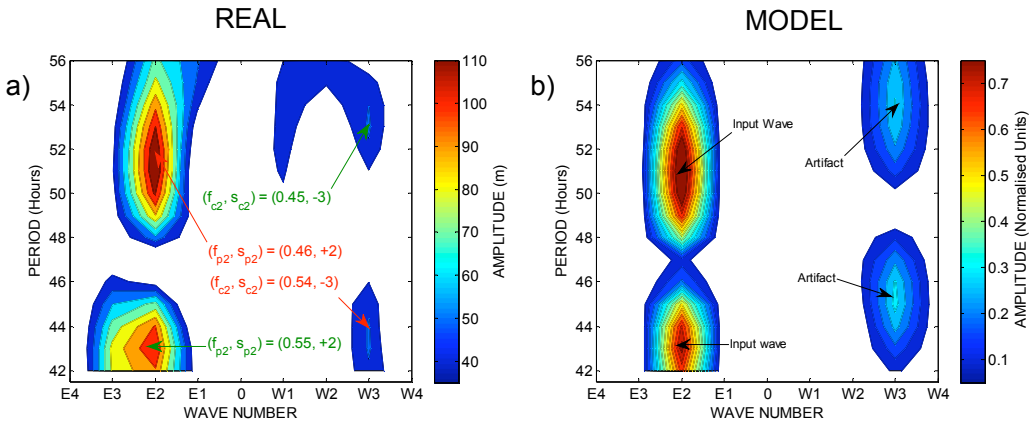


Figure IV.21a: Geopotential height observations of two E2 two-day waves occurring at the same time as two W3 quasi-2-day wave from 1st – 20th December 2004 at a height of about 90 km and a latitude of 68°N. b) Model results from the least-squares fitting with only a W3 wave present in the model input data.

Figure IV.21b again shows model results, however this time, the only waves present in the data are two E2 two-day waves, of slightly different period and amplitude. The two observed W3 waves were not in the model input data and are therefore a result of aliasing.

Using the method described above in section IV.4.1, it can be shown that in our observations up to about 36% of the amplitude of the E2 two-day wave could be aliased into the W3 two-day wave at 68°N, and *vice versa*. This means that during the winter, when the E2 amplitude is actually more than 200% that of the W3 (Figure IV.12b and c), it is not likely that the E2 signal contains a significant aliased component

of the W3. Further, the observation that the E2 obeys Charney-Drazin theorem is also indicative that it is a real wave and not a result of aliasing. We therefore conclude for these two reasons that the observed E2 two-day wave does not contain a significant aliased component – at least in winter. However, we note that in winter 2004/05 and 2006/07 the W3 amplitudes are about one-third those of the E2 and so we cannot be confident that the W3 signal is not entirely caused by aliasing from the E2.

Two mechanisms have been proposed for the excitation of the E2 two-day wave in the stratosphere and/or mesosphere. Firstly, Hartmann (1983), used model studies to examine barotropically unstable flows in the wintertime stratosphere similar to those of the polar night jet.

Two mechanisms have been proposed for the excitation of the E2 two-day wave in the stratosphere and/or mesosphere. Firstly, the repeatability of the E2 wave from year to year suggests that there is a relatively constant forcing mechanism to generate the wave. A relatively consistent source of wave activity in the stratosphere could come from shear instabilities in the polar night jet. These waves are generated by a release of energy caused by baroclinic and barotropic instabilities in the polar night jet. These instabilities are more likely to occur at times of strong zonal winds inside the jet. Baroclinic instabilities are associated with vertical wind shears and barotropic with horizontal wind shears.

Hartmann (1983), used model studies to examine barotropically unstable flows in the wintertime stratosphere similar to those of the polar night jet. The analysis produced normal wave modes, associated with regions of negative potential vorticity gradient on the poleward side of the polar night jet. This instability was strongest for wavenumbers 1 and 2. The wavenumber 1 instability was found to have a 4-day period. Hartmann extended an argument of Frederikson (1982) to suggest that the four-day, wavenumber 1 instability will favour the growth of other instabilities of similar zonal phase speeds. Hartmann noted that an E2 two-day wave will have the same zonal phase speed as an E1 four-day wave and so will promote the growth of the E2 wave. Prata (1984) and Lait and Stanford (1988) presented observational evidence of these four- and two-day waves in the winter stratosphere. This could also go on to suggest an origin for the observed E1 planetary wave, which is likely to have a much higher period, as seen in Figure IV.13c. Baumgaertner et al., (2008) in observational studies, suggested that both barotropic and baroclinic instabilities can contribute to the generation of waves of this type and found that horizontal shear is stronger on the poleward side and vertical shear was stronger on the equatorward side of the polar night jet.

Secondly, Palo *et al*. (2007) reported observations of an E2 two-day wave in the summer mesosphere, which they attributed to a non-linear coupling between the W3 two-day wave and a migrating (i.e., westward 1) diurnal tide.

Our observations suggest that the polar E2 wave is strongly suppressed during major stratospheric warmings (see Figures IV.14 and IV.18). This suppression probably has two explanations. The first is that the winds in the stratosphere do not favour its propagation, as discussed in section IV.4. The second is that an excitation of the wave itself by regions of negative potential vorticity on the poleward side of the stratospheric polar vortex is no longer possible as the polar vortex breaks up. We therefore suggest that the wintertime polar two-day wave is an E2 wave which has propagated upwards from the stratosphere. The amplitudes maximise around the stratopause suggesting that the wave may be evanescent above this height.

IV.6 Arctic Planetary Wave Conclusions

Observations of mesospheric winds over Esrange (68°N) reveal planetary waves with a period near 2 days in summer and in winter. The wintertime two-day wave is much stronger than that observed at mid-latitudes in winter and its structure and origin have been investigated.

Observations made with the Aura MLS instrument reveal the winter two-day wave to be an eastward-propagating zonal wavenumber 2 planetary wave. Although there is a W3 wave present in the winter, aliasing considerations suggest this is likely to be a product of aliasing from the E2 wave rather than a genuine oscillation. The E2 appears to be small in the summertime, which instead is dominated by the classical W3 structure. Because the wintertime W3 wave has negligible amplitude, this suggests that the wintertime polar E2 wave is not generated by non-linear interaction between the W3 and the migrating diurnal tide at mesospheric heights, but rather has been generated in the stratosphere and has ascended to the mesosphere. This interpretation is reinforced by the observation of large E2 amplitudes around the polar winter stratopause. The E1 and E2 waves appear to be the mesospheric manifestation of the four-day E1 and two-day E2 wave observed in the stratosphere by Venne and Stanford (1982), Hartmann (1983), Prata (1984), Lait and Stanford (1988) and Manney *et al*. (1988), the generation of which is attributed to instabilities in the polar night jet. However, further investigation is required to confirm the E1 is actually a four-day oscillation.

We therefore suggest that although “two–day” wave activity is observed at polar latitudes in both summer and winter, the explanation for the summer and winter waves are quite different. The summer two-day wave appears to have the expected W3 (plus W2 and W4) wave structure, whereas the polar winter two-day wave is largely or entirely an E2 wave originating in the stratosphere.

Note that the work presented in this Chapter was published as: Sandford *et al.*, 2008.

Chapter V

V. The Lunar Tide in the Arctic and Mid-latitude MLT Region

The Lunar tide was the first of all large scale fluid motions to be observed, with the tides from the oceans rising and falling approximately twice daily. The connection of these tides with the Moon is thought to have been first recognised and recorded by a famous explorer/mariner of Marseilles called Pytheas in 320 B.C. He observed the tides as he circumnavigated the British Isles and found that there were considerable tidal ranges on the coastlines. He also noted that the times of high water occurred at the same time as the Moon's transit across the meridian. The second high tide, however, when the Moon was on the opposite side of the Earth (out of sight) was of great mystery. An explanation of the tides was first published in Newton's *Principia Mathematica* in 1687. It explained that the tides were a consequence of the solar and lunar gravitational forces. He showed that the solar tide modulated the lunar tide to give the spring tide near new Moon and full Moon and neap tides during half Moon (Chapman and Lindzen, 1970).

Newton explained that the reason for tides seen on both sides of the Earth is that when the Moon is overhead, it has a large gravitational pull on the Earth below, giving the first high tide. At the same time, the opposite side of the Earth feels less gravitational attraction as it is further away from the Moon. However, it also has a large amount of centrifugal force pushing outward as the Earth and Moon rotate about their common centre of mass, which is just below the surface of the Earth underneath the Moon. Therefore the second high tide is created.

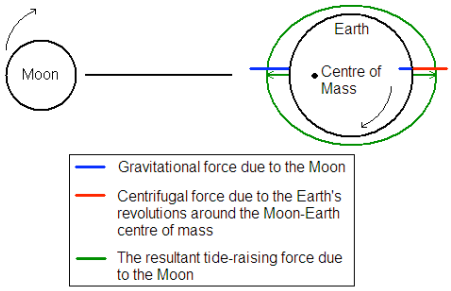


Figure V.1: A north-south cross-section of the Earth-Moon system. Green ellipse represents the tidal forcing on the Earth due to the Moon from the gravitational and centrifugal forces.

In the atmosphere, the lunar tide was also the first to be attempted to be measured by Laplace, who theorised that it could be measured directly at ground level in a large number of hourly pressure measurements. These measurements, however, contained a much larger diurnal and semi-diurnal component than those of the solar tides.

Since those days, the dynamics of the middle atmosphere have been found to be dominated by waves and tides, most of which propagate to this region from sources in the lower heights. Solar thermal excitation produces the well-known 24-, 12- and 8-hour tides that are conspicuous features of the middle atmosphere. The Sun's gravitational field plays only a minor role in the generation of these tides.

In addition to the solar tides, there are also much smaller amplitude tides observed in this part of the atmosphere that have periods related to the apparent motion of the Moon around the Earth. It is believed that the lunar atmospheric tides are generated primarily as a result of the Moon's gravitational attraction on the lower, denser regions of the atmosphere (Stening and Vincent, 1989). However, the vertical motion of the oceans at the lower boundary of the atmosphere may also contribute to the generation of lunar tides (Stening and Jacobi, 2001). This forcing is well specified and virtually constant.

A fundamental difference between tides excited by the Moon's gravitational field and those excited by solar heating is that in excitation of the lunar tides the forcing can be precisely specified, whereas, in the solar case, the tides depend on time-varying atmospheric concentrations and spatial distributions of water vapour, ozone, and molecular oxygen at various heights in the atmosphere. This results in a corresponding variation in the tidal forcing and contributes to the high degree of tidal variability in the observed region.

The solar tides have been extensively studied throughout the atmosphere and especially in the MLT region, where the tides are predicted to maximise. The lunar

tide, however, has been relatively forgotten, primarily due to its amplitude being much smaller than that of the solar tides. Even though it is small in amplitude, it contains a large amount of information, which is not found by studying the solar tides alone. Because the forcing is precisely known, it can be modelled much more accurately than tides whose source regions vary greatly over time and space. This is beneficial for testing of theoretical models of tidal propagation and dissipation in the atmosphere. For example: any variability observed in the lunar tide found in the MLT-region will be primarily due to changes in the propagation characteristics of the tide as it ascends through the atmosphere from its lower source region.

A comprehensive treatment of lunar tides in the atmosphere is given by Chapman and Lindzen (1970). In the Doodson analysis (Doodson, 1922), these are predicted to occur as more than 30 individual components which make up the entire lunar forcing on the atmosphere. Most of these tidal components however, have insignificant amplitudes. It is predicted that the largest amplitudes of these lunar tides will be "migrating" (i.e., Moon following) and have periods relating to the period between lunar transits, 24.84 hours. Of these modes the 12.420-hour (1.9323 cycles per day), is the lunar semi-diurnal M_2 tide or half the lunar transit period. This is predicted to be the largest amplitude for the reason explained above. The second largest is thought to be the principal lunar declinational O_1 tide (Winch and Cunningham, 1972). The O_1 has a period of 25.82 hours and has been observed in geomagnetic data, where it typically has amplitudes of ~ 40 % of the M_2 tide (Stening and Vincent, 1989). The tide predicted to be third largest in amplitude is the N_2 mode. This tide has a period 12.658-hour and therefore occurs at a very similar frequency to that of the M_2 tide. The N_2 arises from the elliptical orbit of the Moon, which varies the lunar distance from the Earth over a period of a month. It has been suggested that this tide has amplitudes of ~ 19 % of the M_2 tide (Schlapp *et al*., 1996).

The other modes also arise because of the complexity in the Moon's orbit. For example, the eccentricity of the Moon's orbit around the Earth gives rise to elliptic tides (i.e., the N_2); the Moon's orbit is tilted with respect to the equatorial plane and this gives rise to declination tides; changes in the eccentricity of the Moon's orbit due to the gravitational attraction of the Sun is able to give rise to tides known as the evectional tides; similarly, changes due to the component of the Sun's gravitational field tangential to the lunar orbit gives rise to variational tides and also the ascending node of the Moon's orbit regresses completely around the ecliptic once every 18.613 years and this gives rise to nodal tides (Winch and Cunningham, 1972). However, these changes in the Moon's orbit are quite small on a day-to-day basis, and therefore they alter the

overall lunar tides by very little. To have a comprehensive treatment of the lunar tide all of these modes would need to be taken into account.

The lunar tide as predicted by classical theory is extremely sensitive to details of the assumed vertical distribution of mean temperature. This suggests that it is subject to interference effects due to multiple reflections caused by variations in the basic temperature (Vial and Forbes, 1994). Miyahara (1975) also showed that lunar tidal amplitudes were generally larger in eastward rather than westward winds. Vial and Forbes (1994) were able to show that using a forcing in their model which was independent of month; they were still able to produce the major seasonal cycles observed in the lunar M_2 tide. This was achieved by using realistic mean wind and temperature profiles.

For the Lunar M_2 tide, the Hough mode solutions have been known for some time. There is a substantial similarity between the Hough functions and equivalent depths of the lunar and solar semi-diurnal oscillations. Most of the excitation goes into the $n = 2$ mode as with the solar S_2 tide (See Chapter II, section II.4.4.1). As there is similarity between the tides, the differences in the solar S_2 and lunar M_2 tides can only arise from the difference in the excitation. For the lunar tide the gravitational excitation enters only through the lower boundary and thus is constrained to a single level (i.e. at the ground). The excitation for the solar thermal tide is, however, distributed throughout a great depth of the atmosphere (Chapman and Lindzen, 1970).

Vial and Forbes (1994) theoretically calculated the lunar tide using Hough Functions and their results showed that strong anti-symmetric tides occur in September and symmetric in December. Observations have shown some differences, these observations have been made by observing the Lunar M_2 tidal phases over two stations at similar longitudes and similar latitudes but in opposite hemispheres. If the zonal phases are within ± 1 lunar hours of each other and the meridional phases are in anti-phase (~ 6 ± lunar hours difference), this would generally indicate the dominance of symmetric modes. However, if the zonal tidal phases were in anti-phase and the meridional phases in-phase then this would indicate the dominance of anti-symmetric modes. This assumes that the dominating mode is not of a higher order than the (2, 4) mode (Stening *et al.*, 1994). Stening *et al.* (1994) obtained the predominance of symmetric tides during solstice months and anti-symmetric tides in equinoctial months. Stening *et al.* (1999) found that the anti-symmetric (2, 3) mode is larger in September but that the symmetric (2, 2) mode is dominant in March. Niu *et al.* (2005) found that the symmetric tide dominated in January, September and December, with the anti-symmetric tide occurring in March, April June and July.

For all these variations, it has been suggested that due to the large amounts of interannual variability different modes could dominate in different years. This could be due to the dominating background winds being different from year to year. Stening et al. (1994) and Niu *et al.* (2005) also suggest that there could also be different tidal modes dominating at different heights.

V.1 Preliminary Results

The first question to be addressed is this: *is a lunar tidal signal detectable in the data set at all?* To answer this question, a Lomb-Scargle analysis was applied to the 16-year set of hourly spaced UK meteor wind radar data as described in Chapter III. This radar measures atmospheric winds at a height of approximately 90 km. The Lomb-Scargle periodogram analysis is based on the least-squares fitting of sine waves to the data, this makes it a very good technique to use with dataset with contain a large number of data gaps or unevenly sampled data. The Lomb-Scargle periodogram provides a measure of the power as a function of frequency (Lomb, 1975; Scargle, 1982).

From the frequency spectrum shown in Figure V.2a, it is possible to see some of the main dynamical features of the MLT region. The figure shows very clear responses at periods 24-, 12-, 8- and 6- hours corresponding to the solar diurnal, semi-diurnal, ter-diurnal and quart-diurnal tides respectively. Examples of studies involving these tides are given by Mitchell *et al*. (2002); Younger *et al*. (2002); Smith *et al*. (2004) and Beldon *et al*. (2006). Also, on the very left of the figure at ~ 0.5 cycles per day (CPD), it is possible to see a small enhanced area, which just about reaches above the significance level. This small hump is likely to be caused by quasi-two day wave activity. The quasi-two day wave has been further investigated in Chapter IV. It should also be noted that amplitude suppression may occur due to the extended time interval used, because of phase cancellation resulting from the seasonal changes in the phase of the tides. i.e., the features could have much larger instantaneous amplitudes. Also to note is that the presence of strong, deterministic signals in the time series (in this case the tides themselves) has the effect of making the confidence level somewhat pessimistic.

The figure also shows a clear response at the period of the M_2 lunar tide; this is the small peak to the left of the solar semi-diurnal tide. The other modes of lunar tide, however, do not appear above the significance level in the figure. Figure V.2b shows

an expanded view of the periodogram of Figure V.2a for frequencies near that of the lunar M_2 tide. The figure shows that, in fact, two closely spaced peaks are present near the M_2 frequency and are of approximately equal amplitude. The frequencies of these peaks are ~ 1.9323 and 1.9350 cycles per day. The first of these corresponds exactly to the M_2 lunar frequency (as near as can be determined by the available spectral resolution). We will take this peak to be strong evidence of a lunar tide being present in the dataset. The second peak corresponds to an oscillation which has almost exactly one more complete cycle per year. An explanation for this observation will be considered in the discussion. A similar analysis applied to the Esrange radar dataset yields similar results (not shown).

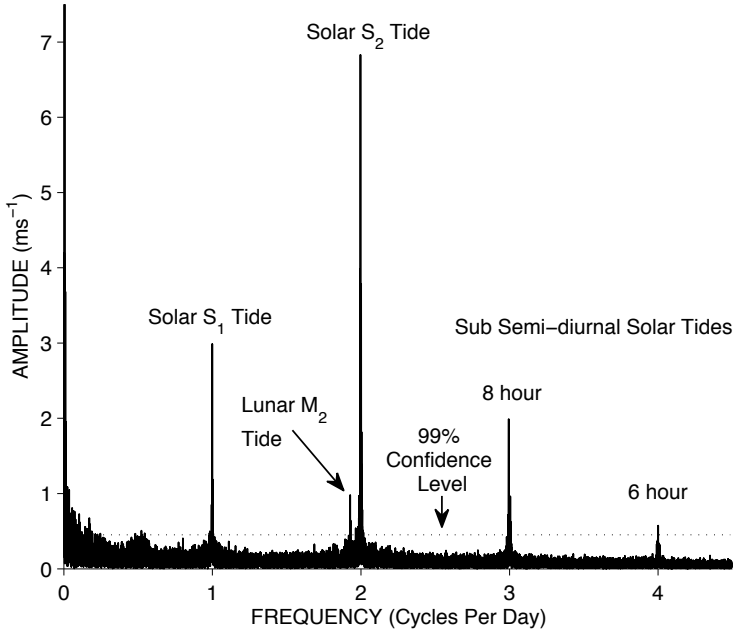


Figure V.2a: Lomb-Scargle analysis of zonal horizontal wind data over an interval of 16 years (1988 - 2004) from the UK meteor radar.

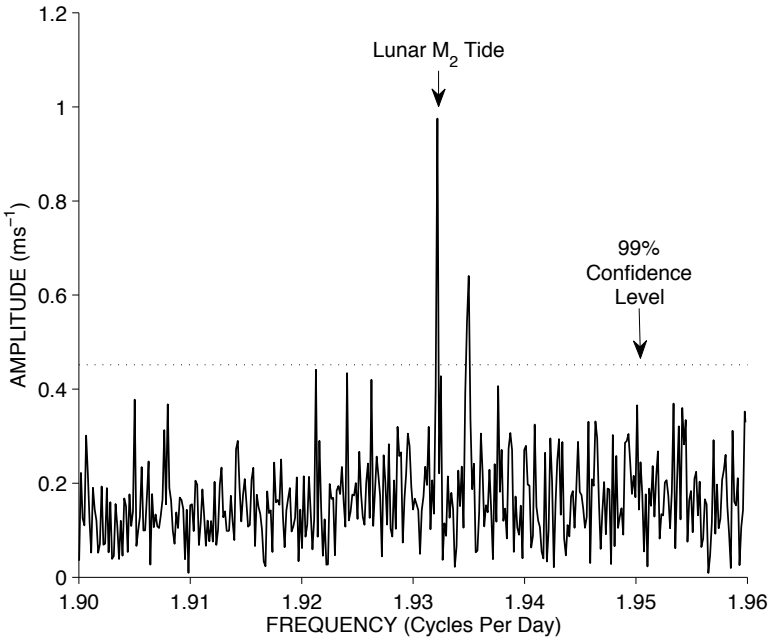


Figure V.2b: A zoomed in view of the Lomb-Scargle periodogram shown in Figure V.2a, shows a peak at the frequency corresponding to the M_2 lunar tide.

For the purpose of this study only the lunar M_2 tide is seen to reach significant amplitude in the data sets used. Therefore, the focus will be given to the 12.420-hour M_2 lunar tide. Observational studies of the tide have, to date, been comparatively limited. This is because it has proved difficult to study due to its small amplitude when compared to the solar tides. The lunar M_2 tide is also the only component of the lunar tide on which other authors generally concentrate. In the MLT region, the M_2 tide is thought to seldom reach amplitudes of more than ~ 10 ms^{-1}.

A second problem in analysis, specifically for the M_2 tide, is its proximity in frequency to the solar S_2 tide, as shown in Figure V.2a. Because their frequencies are so close, there is a need for extended datasets to provide the necessary spectral resolution to distinguish the solar and lunar components.

To resolve the two frequencies, it is required that one of the waves goes through one complete cycle more than the other wave. i.e., in a 24 hour period, the solar S_2 tide goes through 2 complete cycles, whereas the lunar M_2 tide only goes through 1.9323 cycles. This is shown in Figure V.3, assuming the S_2 and M_2 have the same amplitudes.

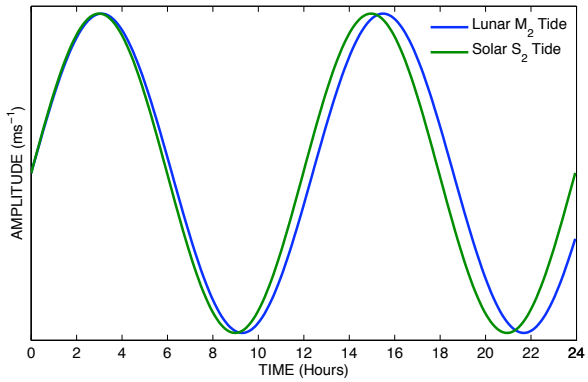


Figure V.3: The first 24 hours of an S_2 and M_2 tide with equivalent amplitude.

For the S_2 tide to travel through one complete cycle compared with the M_2 requires 14.7710 days. Therefore, this requirement means that it is only possible to discriminate between the M_2 and S_2 frequencies in data sets where there are more than about 15 days of measurements.

Despite the difficulty in identifying lunar tidal signatures in the MLT region, a number of studies have detected their presence. These studies have invariably relied on long data sets of up to ~ 13 years duration. Various techniques have been used, including MF radar (e.g., Stening and Vincent, 1989; Schlapp and Harris, 1993 and Stening *et al*., 1987, 1994, 1995, 1997b); MST radar (e.g., Stening *et al*., 1990); meteor wind radar (e.g., Tsuda *et al*., 1981; Stening *et al*., 2003; Niu *et al*., 2005 and Sandford *et al*., 2006, 2007a,b) and LF-D1 (Stening and Jacobi, 2001). These techniques are particularly suitable because they are all capable of yielding continuous time series of many months duration. The above studies have revealed that the horizontal winds associated with the M_2 lunar tide can reach amplitudes as large as ~ 10 ms^{-1} in the MLT and, in general, have revealed lunar tides to have a distinct seasonal cycle, with maximum amplitudes usually occurring around midwinter and autumn in the northern hemisphere. Vertical wavelengths are observed to be quite short, ranging from ~ 17 km to ~ 80 km.

A very limited number of modelling studies have addressed lunar tides in the atmosphere. Vial and Forbes (1994) considered the seasonal variability of the M_2 mode from the ground to a height of 105 km. The model includes forcing by the Moon's gravitational attraction on the atmosphere, as well as a dynamical effect caused by the vertical motion of the lower boundaries (land and oceans) and also forcing due to the load tide, which accounts for the deformation of the solid Earth due to the added weight of water above it (the load tide is approximately out-of-phase with the ocean tide). This model only includes zonal wavenumber, $s = 2$ component. The

model assumes a small perturbation of a shallow, compressible, hydrostatic, perfect gas atmosphere. The equations of momentum, mass continuity, thermal energy and an ideal gas are linearised and solved as in Chapter II. Dissipation processes are also included in the solutions and these are molecular and eddy diffusion, Newtonian cooling and ion drag (Vial and Forbes, 1994). This model is a zonal-mean model showing the latitudinal variations in the M_2 lunar tide. At meteor heights, amplitudes were predicted to reach up to ~ 10 ms^{-1}. Figure V.4 shows a latitude-time cross-section of predicted Lunar M_2 tidal amplitudes from the Vial and Forbes (1994) model at approximately 90 km. Both the zonal and meridional components are very similar, however, the meridional has higher equatorial amplitudes. Figure V.4 shows that the highest concentration of amplitudes occurs in the northern hemisphere between mid to high latitudes during the late summer and autumn months.

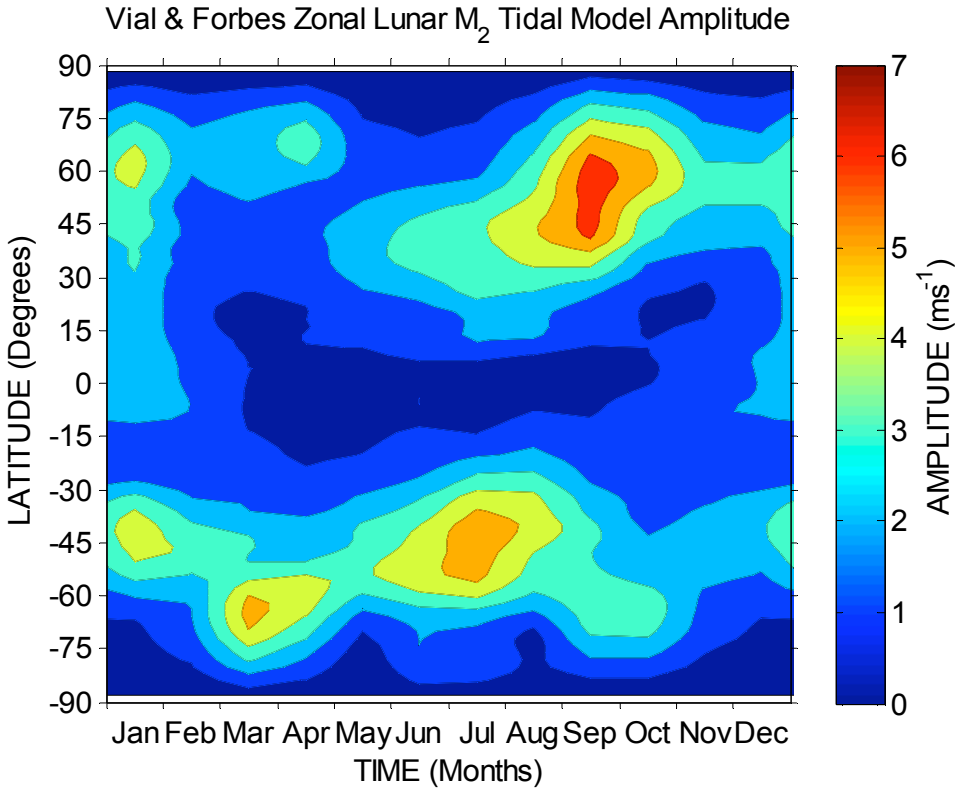


Figure V.4. Latitude-time cross-section of zonal amplitude of the lunar M_2 tide as predicted by the Vial and Forbes (1994) model at a height of about 90 km.

In the work presented in this chapter, observations of the M_2 lunar tide have been made using two meteor radars. One radar is located at middle latitudes, the other in the Arctic. Data from several years are investigated to establish the seasonal variability of the tide and to determine its vertical structure at Arctic latitudes.

V.2 Data Analysis

The data used in this chapter have been collected from two meteor radars - one at Castle Eaton (52.6°N, 2.19°W) in the UK and the other at Esrange (67.9°N, 21.1°E) in Arctic Sweden. The UK radar is a narrow beam system using two orthogonal beams operating at 36.75 MHz. This radar has been in near-continuous operation since January 1988 (although with a one-year gap in 1995-96 when only one beam was used). Horizontal wind components measured in the two beams are combined to give hourly-spaced estimates of the zonal and meridional wind. The data set considered here did not include an estimate of the heights of individual meteors, so the winds are representative of the MLT-region, weighted by the vertical distribution of meteor echoes at the radar's operating frequency. A description of the Castle Eaton radar is given in Chapter III and can also be found in Muller *et al*. (1995) and Beldon *et al*. (2006).

The Esrange radar has been described in Chapter III and also by Mitchell *et al*. (2002). It is a SKiYMET, all-sky system with a transmitter frequency of 32.5 MHz. Continuous data has been recorded from October 1999 with no significant gaps. The radar uses a set of five receiving antennas as an interferometer to determine the meteor trail azimuth and zenith angles, which, in combination with range information, allows the determination of echo height. Routine range resolution is ~ 2 km. If the reader refers to Figure III.10b in Chapter III it is seen that the height distribution of meteors, using a radar of 32.5 MHz is strongly peaked at ~ 90 km with only very weak seasonal variation. Therefore the winds derived from the UK meteor radar, using the slightly higher frequency of 36.75 MHz, are also likely to be centred at ~ 90 km.

To again test that there is a detectable lunar tidal signal in the data sets being used, following on from Figure V.2a and b, a second test is provided by a superposed epoch or "composite lunar day" analysis. Because the lunar tides are "Moon following" rather than "Sun following", they will not repeat from day to day with the same phase as measured in local time. Instead, an analysis must be applied that relates the tidal phase to the position of the Moon in the sky. In this case, zonal and meridional winds were sorted by lunar time. Using this analysis, averaged over extended intervals, any coherent feature which repeats itself with a period equal to exactly n times the lunar transit frequency (where n is an integer) will be evident, whereas any oscillations present with frequencies not related to the lunar-transit frequency will self cancel and so not be evident.

An example is given in Figure V.5 which presents composite lunar day analyses of data recorded over Esrange from the three-month interval December 2000 – February

2001. Figure V.5a presents this analysis applied to the zonal component and Figure V.5b presents it for the meridional component. Note that at each height gate the mean has been subtracted from the time series so that any oscillations present are not lost against large changes with height resulting from the mean flow. The plots are shown in lunar time; one lunar day is defined as 24.84 ± 0.016 hours (24 hours 50 minutes ± 1 minute), the time it takes for the Moon to make one complete orbit of the Earth. One lunar hour is thus ~ 1.0347 solar hours.

In each figure a clear wave-like structure is evident across the entire height range of 70-90 km considered. Two cycles are evident per lunar day, suggesting that this feature has the same period as the lunar semi-diurnal (M_2) tide. The amplitude of the feature seems to be ~ 8 ms^{-1}. Further, a clear phase progression is observed, which in the case of a wave or tidal motion would correspond to an upwardly-propagating wave or tide.

An obvious concern is that spectral leakage may occur from the large-amplitude solar, S_2, semi-diurnal tide to lunar semi-diurnal tidal frequencies, since the two frequencies are very close. To investigate this possibility, an analysis similar to that above was carried out, but using a superposed epoch period of 23.16 hours instead of 24.84 hours. This new period differs from that of the 24.0-hour solar day length by exactly the same amount as does the lunar day length, except that in this case it is shorter rather than longer. Figures V.6a, b show the results of this analysis, presented in a similar format to that of Figures V.5a, b and for the same time period. If it were the case that the wave-like features of Figure V.5a, b were caused by spectral leakage from the S_2 solar tide, then we might expect a similar leakage into this higher-frequency analysis. However, although there is some variability within the composite day, it does not have the clear wave-like features of Figure V.5 and those features that are present are of smaller amplitude (~ 4 ms^{-1}, c.f. ~ 8 ms^{-1}). We take this as further evidence that a significant lunar tide is present and identifiable in the data.

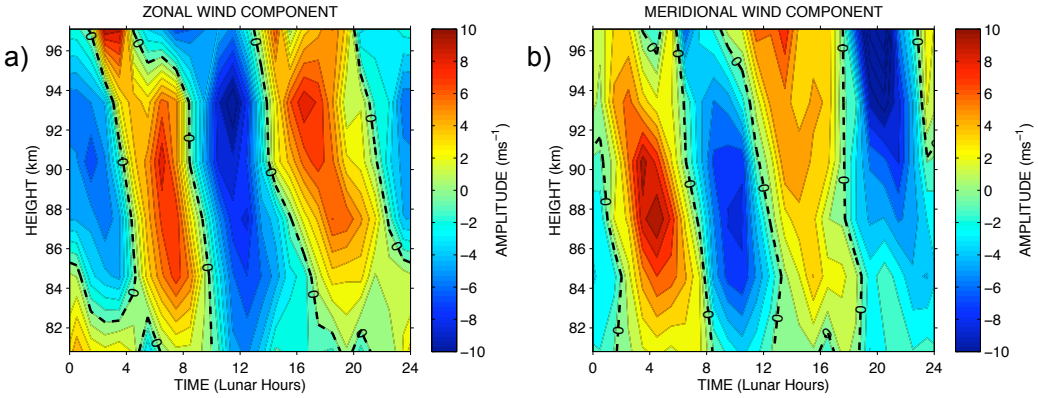


Figure V.5. Composite lunar day of zonal (a) and meridional (b) horizontal wind data for winter (DJF) 2000/2001 over Esrange (1 Lunar day = 24 Lunar hours = 24.84 Solar hours).

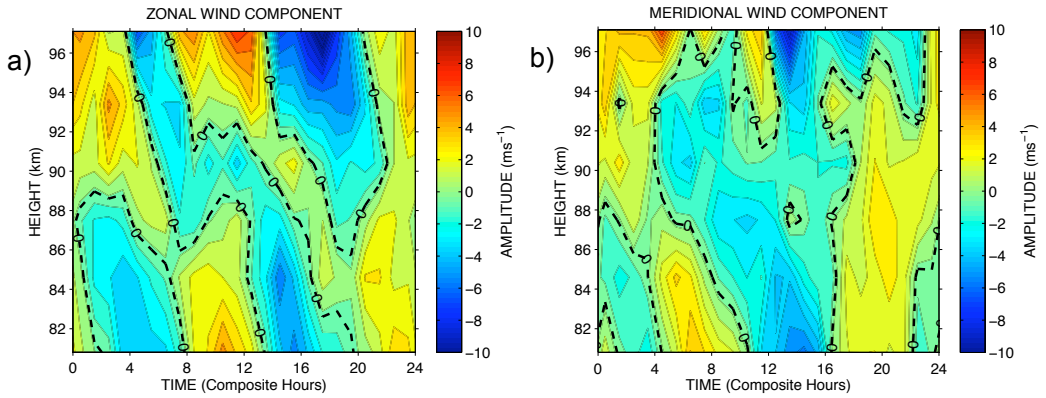


Figure V.6. Composite day of zonal (a) and meridional (b) horizontal wind data for winter (DJF) 2000/2001 over Esrange (1 composite day = 24 composite hours = 23.16 Solar hours).

In summary, a spectral and a composite lunar-day analysis reveal a coherent wave-like feature which has exactly two cycles per lunar day. The feature exhibits a phase gradient corresponding to that expected for an upwardly-propagating wave or tide. A simple analysis based on searching for similar features spaced at an equal frequency from the solar S_2 tidal frequency suggests this feature is not caused by spectral leakage from the solar tide. We therefore conclude that the wavelike features evident in the superposed-epoch analysis of Figure V.5 and the spectral analysis of Figure V.2 are, in fact, the M_2 lunar semi-diurnal tide.

Studies by other authors have investigated a number of different methods of isolating the lunar tide from time series of geophysical data. Stening *et al*. (1997b) compared least-squares fitting of a sum of solar and lunar tides to geophysical time series (Malin and Schlapp, 1980) in comparison to methods based on Fourier analysis (Winch and Cunningham, 1972) and concluded that the least-squares method offers a number of advantages. In particular, each hourly data point is treated separately, so scattered and/or missing data is not a problem. Further, with this method, results can be accurately tuned to the required tidal frequencies.

In the work presented here, the analysis follows the technique of Stening *et al*. (1997b). The time series of zonal, u or meridional, v horizontal winds, are taken to consist of a superposition of a number of tidal oscillations on a background flow. The *a priori* assumption is made that the major tides are sinusoidal oscillations with arbitrary phases (phase being measured from the East) and amplitudes and that these tides have periods of 24-, 12-, 8-, 6-hours and the M_2 period. These correspond to the first four principal solar tides and the M_2 lunar tide, as per the equation:

$$u,v(t) = \overline{u},\overline{v} + \sum A(\omega)\cdot\cos\left[\frac{2\cdot\pi\cdot\omega\cdot t}{24} - \phi(\omega)\right] \qquad \text{V.1}$$

Where $\overline{u}$ or $\overline{v}$ is the mean wind, A is the amplitude of a particular tide, ϕ is its phase, t is the time step in hours and ω is its frequency in cycles per day.

To determine the period of the Lunar M_2 tide we consider the main term in the gravitational potential generating the lunar semi-diurnal tide. This varies as $\cos(2\tau)$, where τ is the "lunar time" and is given by $\tau = t - \nu$. Here t is the local solar time and ν is the lunar age, a cyclic term which is dependant on the phase of the Moon, and where new Moon is signified by $\nu = 0$ (e.g., Stening *et al*., 1987; Stenning and Vincent, 1989). This relates the position of the Moon and Sun with the Earth, so that the lunar tide can be measured in solar hourly spaced data.

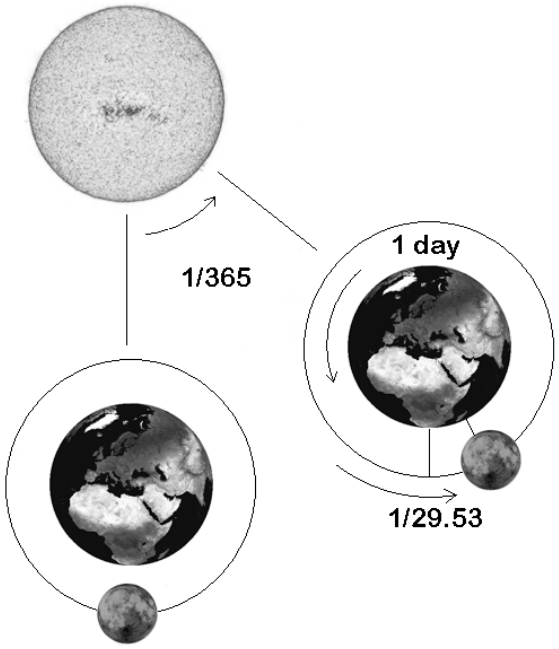


Figure V.7: The movement of the Moon in 24 hours as viewed from above the ecliptic plane (Not to scale).

A key difference between the solar and lunar tides is that the former are Sun following, and so (allowing for seasonal variations) the tide has a phase that is consistent in local time from day to day. In contrast, the lunar tides are Moon following, and so the phase will decrease systematically from day to day when measured in local solar time. Lunar tidal phases must therefore be measured with respect to the position of the Moon. In 24 hours the Earth makes one complete revolution and the Sun is seen to return to the same position in the sky above the Earth as in Figure V.7. The Moon, however, takes 24 hours and an extra ~ 50 minutes, as it rotates in the same direction as the Earth, before it appears in the same position in the sky i.e., ~ 1/29.53 (1/ lunar month) = 0.03386 days = 48.76 minutes.

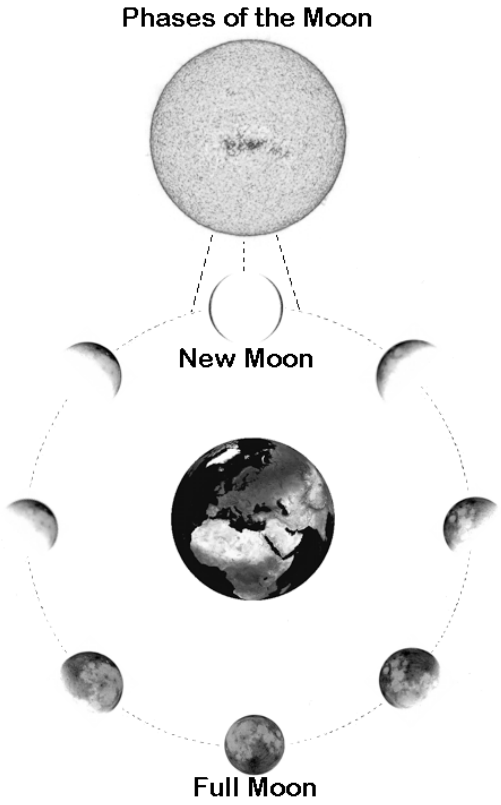


Figure V.8: The phases of the Moon (After Verne, 1865)

A lunar month is the time between successive new Moons. A new Moon is when the Sun, Earth and Moon are aligned with the Moon in between the Sun and Earth, see Figure V.8. This time between new Moons varies very slightly over time as shown in Figure V.9a. It is at its longest when the Moon is moving slowest near apogee and the Earth is moving fastest at perihelion, and *vice versa*. The period of a tide created by the Moon would therefore be 1 – 1/29.53 in cycles per day. This gives the diurnal component of the tide and twice this value gives the semi-diurnal component as shown in Figure V.9 (bottom plot).

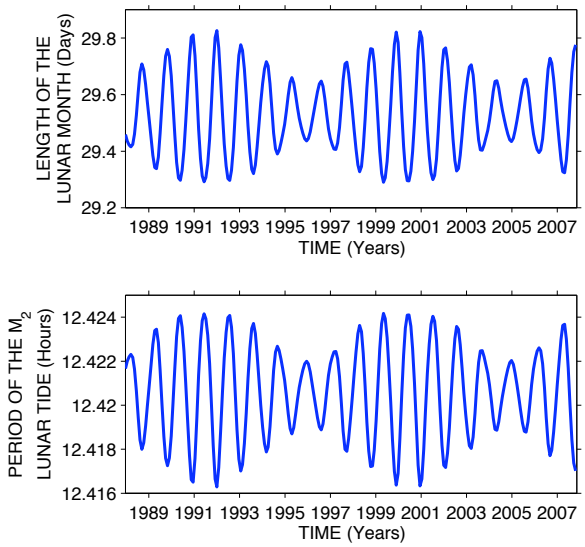


Figure V.9: Top plot indicates the length of the lunar month in days, i.e., the time between new Moons. The bottom plot indicates the period of a semi-diurnal lunar tide observed from a static location on the Earth.

The declining mean orbital eccentricity of the Earth also tends to reduce the range of lunar month variations. The average lunar month (i.e., the mean synodic month) has miniscule long-term change compared to this shorter-term periodic variation (Bromberg, 2007).

In the work presented here, the fitting period will not be exactly constant but it varies over the period 12.420 ± 0.0043 hours, this includes the M_2 tide and a very small amplitude-long period lunar tidal oscillation (Winch, 1981). On a monthly-mean basis, this does not vary significantly from the M_2 tide, therefore, the tide calculated herein, will be referred to as the M_2 lunar tide. The time series of horizontal winds are analysed in monthly-mean segments to determine the amplitude and phase of each of the tidal components (S_1, S_2, S_3, S_4 and M_2) plus a mean wind for that particular time series segment. Each segment is greater than 15 days in length as investigation into the sampling theory above shows that this is required in order to resolve the lunar M_2 tide from the solar S_2 tide. This approach also has the additional advantage of allowing easy comparison to other observations and to the monthly-mean results from the model of Vial and Forbes (1994).

An example of the response to this fitting technique using model tidal data is shown in Figure V.10. This fitting uses a constant amplitude wave of varying frequency and fixed phase. It is fitted using the least squares lunar tidal fit, as described above. As can be seen from Figure V.10, the maximum frequency is at the lunar M_2 tidal frequency of 1.9323 cycles per day. The figure also shows that there is a minimum at

the solar S_2 tidal frequency of 2 cycles per day, which helps in avoiding spectral leakage from the sidebands.

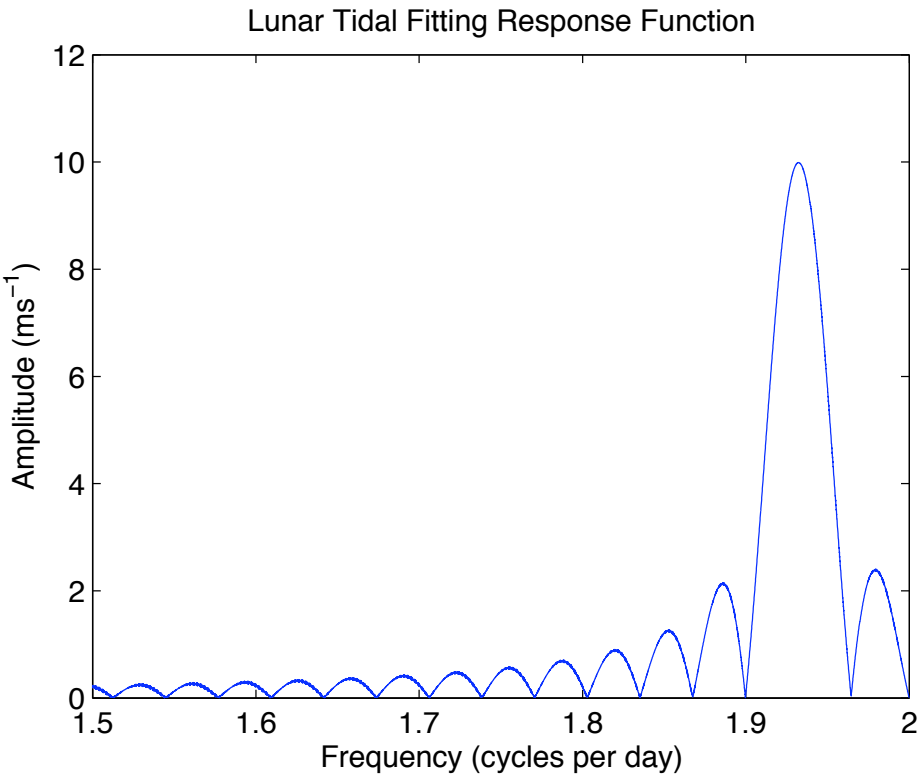


Figure V.10: The response to a lunar tidal fit on data, which has a fixed amplitude and phase, but varies in frequency.

In the case of the Esrange radar, the analysis was applied to winds calculated in six height gates. The height gates used have depths of 5, 3, 3, 3, 3 and 5 km (from 78 – 100 km). Because of the uneven vertical distribution of meteor echoes, which is strongly peaked at ~ 90 km, the winds calculated for each height gate are assigned to the average meteor echo height for that gate. This leads to the winds being calculated at representative heights of 80.8, 84.6, 87.5, 90.4, 93.4 and 97.1 km. Note that each height gate is independent of the others.

Note that because of the approximate semi-diurnal nature of the lunar tide, all phases can be arbitrarily shifted by 12 lunar hours.

V.3 Results

V.3.1 Mid-Latitude Seasonal Behaviour

Monthly-means of amplitude and phase were calculated for tides of period 24-, 12-, 8-, 6- and 12.420- hours in all of the months available in the 16 years of the UK data set. The monthly-means of amplitude and phase for the M_2 lunar tide for each month from successive years were then vector averaged to reveal the seasonal variability. Figure V.11a presents the vector-averaged monthly-mean amplitudes for the zonal and meridional components. The error bars indicate the standard error on the mean. Also indicated are the monthly-mean amplitudes from the Vial and Forbes (1994) model for a height of 93 km (selected as being the height closest to that of the meteor-radar observations). The figure shows that the observed zonal and meridional amplitudes are generally quite close. There is also a clear seasonal cycle apparent in which largest amplitudes are observed in winter (January - February) and autumn (September – October). Amplitudes reach an extended minimum throughout the spring and summer months (March – August) and there is a distinct drop in amplitudes evident in November, following the autumnal maximum and before the amplitudes rise again in winter.

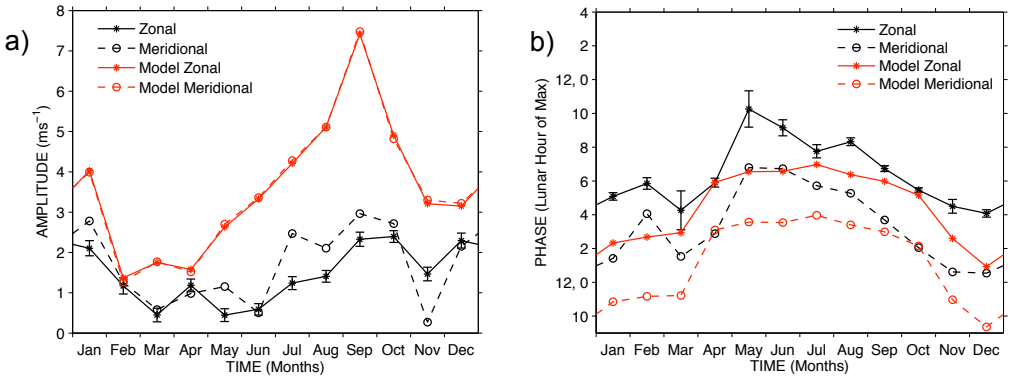


Figure V.11a, b. Monthly-mean a) amplitudes and b) phases of the Lunar M_2 tide calculated from UK data (1988 – 2004) vector averaged to produce the seasonal variation indicated in black. Red lines indicate model amplitude results from Vial and Forbes (1994) model for 93 km. Solid lines represent the zonal component and dashed the meridional. Error bars indicate the standard error on the mean.

A comparison with the model shows that it captures the observed general seasonal behaviour quite well, although the model amplitudes are significantly higher – maximising at some ~ 8 ms^{-1} in September compared to the observed ~ 2 – 3 ms^{-1}. We should note, however, that at meteor heights the model amplitudes increase rapidly with height and so small height differences result in better or worse agreement.

Figure V.11b presents the observed vector-averaged monthly-mean phases for the zonal and meridional components over the 16 years of data. A systematic phase difference of ~ 3.034 lunar hours is seen between the zonal and meridional components when the amplitude is high. This corresponds to a phase angle of ~ 91° as the full cycle for the lunar tide is 12 lunar hours. A systematic phase difference of ~ 2.68 lunar hours is seen if all months are considered, corresponding to ~ 80.2°. These values, in combination with the approximately equal zonal and meridional amplitudes, correspond to an approximately circularly-polarised oscillation in which the wind vector rotates in a clockwise manner. i.e., the tide maximises in the northward direction first, then ~ 3 lunar hours later it maximises in the eastward direction, then ~ 3 lunar hours later maximises in the southward direction and so on. This agrees with tidal theory in the northern hemisphere for a rotating wind vector, i.e., the Coriolis force acts to the right in the northern hemisphere, producing a clockwise rotation.

The phases themselves show a clear seasonal cycle. The phases have their smallest values in winter and early spring (December – March); they advance rapidly by some ~ 6 lunar hours in late-spring to early summer (April – May) and then decline at a fairly constant rate throughout the summer and autumn back to the winter values. Also indicated in Figure V.11b are the corresponding phases from the Vial and Forbes model for a model height of 93 km. In comparison to the observations, it can be seen that the model reproduces the seasonal behaviour well, but that there is a systematic phase difference in which the model phases lead the observed phases by ~ 2.12 lunar hours.

V.3.2 Mid-Latitude Inter-Annual Behaviour

Finally, we should note that the figures indicate that the M_2 lunar tide exhibits a high degree of inter-annual variability in both amplitude and phase. The standard deviations of the amplitudes and phases are significant, being ~ 4 times larger than the standard errors indicated. The phase fluctuations mean that there is some degree of cancellation in the vector-averaging of the amplitudes. It was found that geometric averaging of the amplitudes resulted in averaged monthly means some ~ 72 % larger than those seen by vector averaging.

To further investigate this inter-annual variability, Figure V.12 presents the annual geometric mean of the tidal amplitude (i.e., the phase is not taken into account) for the interval 1988 – 2005, separated into the four seasons. To simplify the figures, the zonal and meridional amplitudes are averaged together. The gap in the data from

March 1995 to December 1996 when only one beam of the radar was operating and is filled using measurements made in that beam only. The figures show that there is considerable year-to-year variability in seasonal-mean amplitudes, which vary by more than a factor of two from year to year.

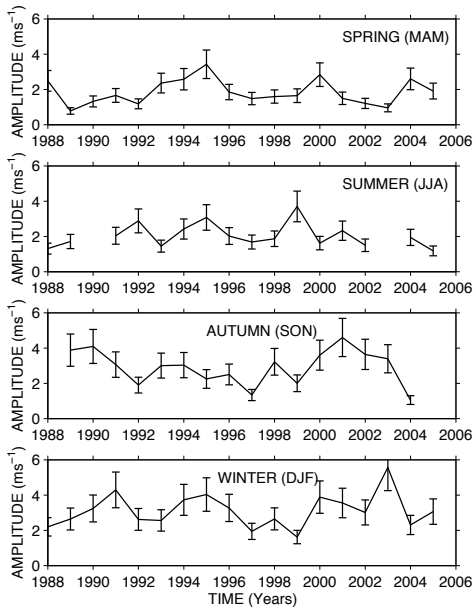


Figure V.12. Year to year variability of the Lunar M_2 tidal amplitude calculated using monthly – mean amplitudes of the zonal UK data (1988 – 2004) averaged into seasons. Error bars indicate the standard error on the mean.

V.3.3 Arctic Seasonal Behaviour

A similar analysis to that performed with the UK data was applied to data recorded by the Esrange meteor radar – although in this case six independent height gates were available between 80 and 100 km, allowing an investigation of the vertical structure of the tide. Firstly the six-height gates are averaged to give a similar plot to that of Figure V.11. Figure V.13a presents the vector-averaged monthly-mean amplitudes for the zonal and meridional components. The error bars indicate the standard error on the mean. Also indicated are the monthly-mean amplitudes from the Vial and Forbes (1994) model averaged over the same height range.

The figure shows that the observed zonal and meridional amplitudes are generally quite close as with the UK results. The clear seasonal cycle is also apparent, in which largest amplitudes are observed in winter (January - February) and autumn (September – October). Amplitudes reach an extended minimum throughout the spring

and summer months (March – August) and there is again the distinct drop in amplitudes evident in November following the autumnal maximum. Comparison of the observations with the model shows that again it has captured the observed seasonal behaviour quite well, although during the winter, amplitudes tends to be higher and during the secondary maximum, the observed amplitude tend to be lower. In fact, the model predicts that the September maximum will actually be larger than the winter time maximum. The gradual rise seen in the model from the spring minimum toward the September maximum is also not observed.

Figure V.13b presents the observed vector-averaged monthly-mean phases for the zonal and meridional components over the seven and a half years of Esrange data. The systematic phase difference between the zonal and meridional components is ~ 3.34 lunar hours which corresponds to a phase angle of ~ 100.2°. Combined with the approximately equal zonal and meridional amplitudes, it corresponds to an approximately circularly-polarised oscillation in which the wind vector again rotates in a clockwise manner (i.e., the tide is in quadrature).

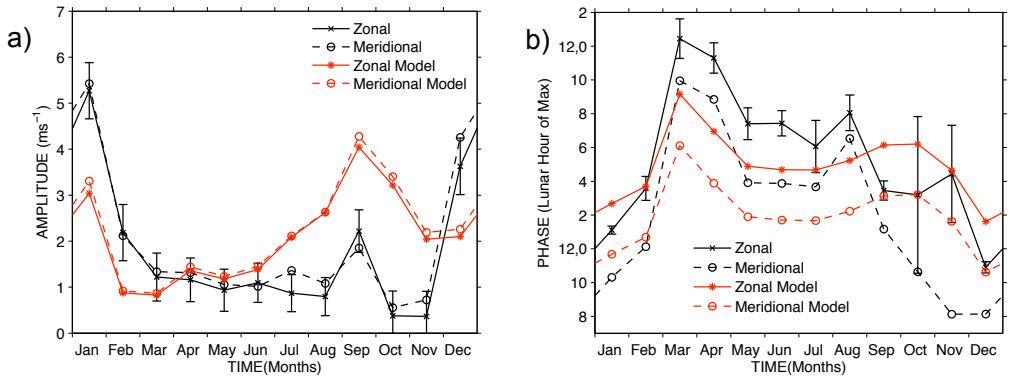


Figure V.13a, b: Monthly-mean a) amplitudes and b) phases of the Lunar M_2 tide calculated from Esrange data (1999 – 2007) in the same format as Figure V.11. The amplitudes are vector averaged to produce the seasonal variation indicated in black. Red lines indicate model amplitude results from Vial and Forbes (1994) model for 93 km. Solid lines represent the zonal component and dashed the meridional. Error bars indicate the standard error on the mean.

To investigate the vertical structure of the tide, Figures V.14a, b present contours of the zonal and meridional amplitude of the lunar M_2 tide. As before, the data have been vector averaged to produce a composite year from the seven and a half years of data available (October 1999 –August 2007).

Both sets of contours reveal a generally similar and well-defined seasonal behaviour in which the largest amplitudes of ~ 8 ms^{-1} are reached in winter (December – February) with the secondary maximum around August-September. Amplitudes are generally

small and relatively constant (< 2 ms^{-1}) during an extended minimum, occurring throughout the spring and summer months (March – July). Except at the lowest heights, the ratio of maximum to minimum monthly-mean amplitudes is larger than over the UK. Amplitudes rise with increasing height in all months (except in September where a localised amplitude maximum occurs near 87 km). The small reduction in amplitudes evident in November in the UK data (Figure V.11a) also has a counterpart in the amplitude minimum observed during October-November in the Arctic.

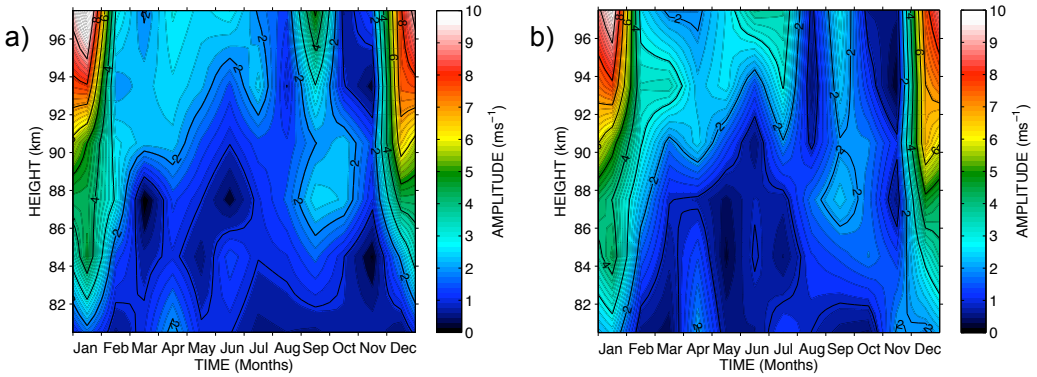


Figure V.14a, b. Contour plot of the Lunar M_2 tidal amplitude calculated from monthly-mean amplitudes (vector averaged to give the seasonal variation) of a) zonal and b) meridional wind data over the period 1999 – 2007 from the Esrange meteor radar.

To compare this behaviour with that of the solar S_2 semi-diurnal tide, the same analysis was applied. Figures V.15a, b present amplitude contours of the a) zonal and b) meridional components of the solar S_2 tide. As before, the data have been vector averaged to produce a composite year from the seven and a half years of data (October 1999 – August 2007).

Both sets of contours in Figure V.15, reveal very similar behaviour to that seen in Figure V.16. There is a well-defined seasonal behaviour in which the largest amplitudes in winter (December – February) reaches amplitudes up to 35 ms^{-1} and the secondary maximum around August-September. The extended minimum throughout the spring and summer months (March – July) is also seen as well as the reduction in amplitudes evident in November. The amplitudes of the solar S_2 tide are much larger than those of the lunar M_2 tide, by a factor of about three. The amplitudes also tend to rise with height as seen with the lunar tide and the zonal and meridional components are very similar in amplitude.

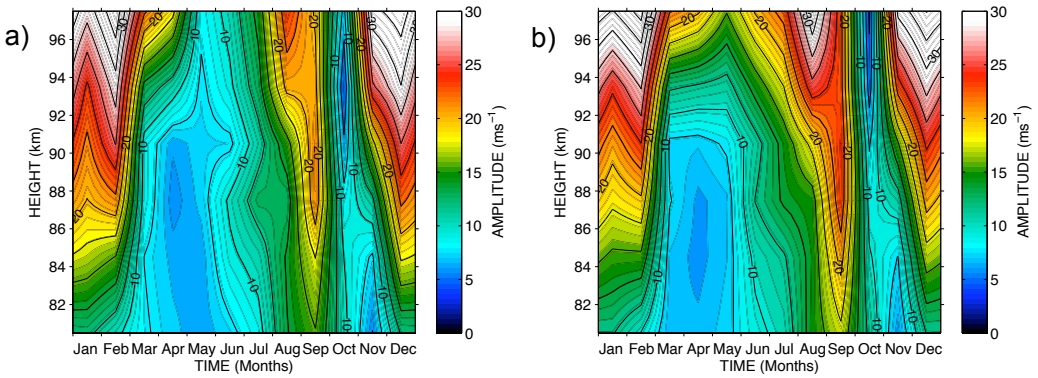


Figure V.15a, b: Contour plot of the Solar S_2 tidal amplitude calculated from monthly-mean amplitudes (vector averaged to give the seasonal variation) of a) zonal and b) meridional wind data over the period 1999 – 2007 from the Esrange meteor radar.

The lunar M_2 tidal results from the Arctic are now compared with the Vial and Forbes (1994) model. Figure V.16 presents height profiles of amplitude for each month in which the Esrange observations are plotted along with the predictions of the model. Both the zonal and meridional components are plotted. In most months the model amplitudes agree well with the observations (i.e., within a few ms^{-1}). However, from the autumnal equinox to early winter (September to November) the observed amplitudes in the upper heights are systematically smaller than those of the model. For instance, in October the model predicts a rapid growth of amplitude with height to amplitudes of greater than 5 ms^{-1} at heights above 90 km. The observations reveal very limited growth of amplitude, with amplitudes less than 2 ms^{-1}. In contrast, during December-January, the observations reveal amplitudes that are significantly larger than predicted. For instance, in December, the rapid growth of amplitude with height occurs rather lower than in the model. The spring to early summer observations tend to match very well at generally all heights. In summary, the observations reveal smaller amplitudes than predicted in autumn and early winter, and larger amplitudes than predicted around the winter solstice.

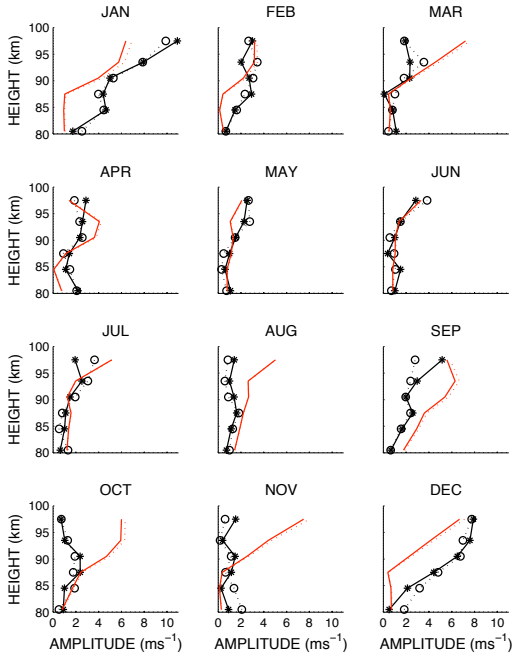


Figure V.16. Monthly amplitude plots of the Lunar M_2 tide. Black symbols indicate tidal amplitudes calculated from the Esrange meteor radar over the period 1999 – 2007 and the red lines indicate results from the Vial and Forbes (1994) model. Solid lines show the zonal (stars) component and dashed show the meridional (circles).

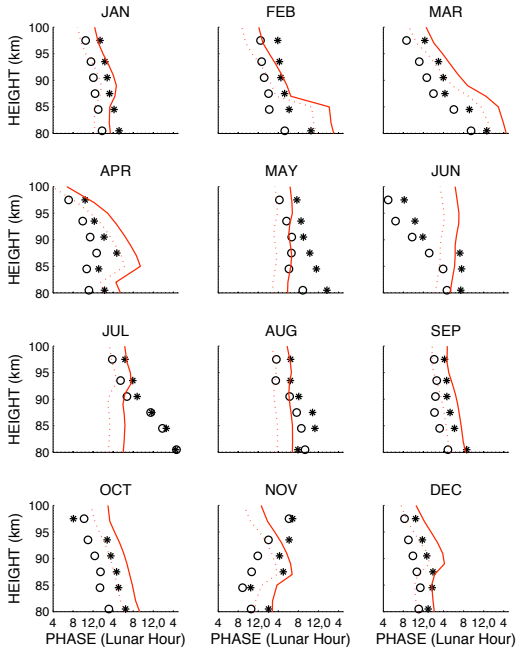


Figure V.17. Monthly phase plots of the Lunar M_2 tide. Black symbols indicate tidal phases calculated from the Esrange meteor radar over the period 1999 – 2007 and the red lines indicate results from the Vial and Forbes (1994) model. Solid lines show the zonal (stars) component and dashed show the meridional (circles).

Figure V.17 presents a similar analysis applied to the phases. Again, both the zonal and meridional components of the observed and model tides are presented as vector averages of data over the seven and a half years observed. The phase indicates the hour of maximum wind in the eastward direction for the zonal component and northward for the meridional. Generally, the phase behaviour shows the meridional component leading the zonal by ~ 3 lunar hours, corresponding to the phase quadrature at all heights, with a wind vector which rotates clockwise with time (when viewed from above). It is immediately noticeable that in general the phase decreases with increasing height. From tidal theory, this indicates the energy of the waves is upwardly-propagating. Error bars are not shown in the figure for reasons of clarity; however, the mean uncertainty (error on the mean) in phase over the entire ensemble of monthly data is 0.9 lunar hours. The uncertainty in phase is largest when the amplitudes are smallest. The largest uncertainties occur in October, when the average error bar has a value of ~ 3.3 lunar hours. The smallest uncertainties occur in January, when the average error bar has a value of ~ 0.3 lunar hours.

It is noticeable that when the amplitudes are large (e.g., in September, December and January) the phase progression of the tide with height is very smooth and also the observations follow the model very well. This regular progression of phase and consistent phase quadrature over the six independent height gates observed provides further confidence that a clear tidal signature has been identified.

The variation of tidal phase with height allows calculation of the vertical wavelength for each month. Table V.1 lists the observed vertical wavelengths and those from the model of Vial and Forbes (1994). The vertical wavelengths are calculated using a least-squares fit to the phase versus height data and also with the model results versus height. An average of the zonal and meridional components is shown and the errors indicate the quality of the fit. Phases with the largest errors (i.e., those with the smallest amplitudes) have been excluded from the vertical wavelength determination, but are still shown in Figures V.16 and V.17. In general, the observations match reasonably with the modelled results, with larger values in winter and summer, and smaller values in spring and autumn. Towards the end of spring and summer, the modelled phase sometimes *increases* with altitude, i.e., indicating a downward propagation. This results in a negative vertical wavelength.

Table V.1. Monthly-mean vertical wavelengths observed over Esrange (68°N, 21°E) from October 1999 to August 2007 and those from the model of Vial and Forbes (1994). The wavelengths are calculated over the height range 80 – 100 km and zonal and meridional values averaged. Errors indicate the error on the fit to the gradient.

Month	Measured Vertical Wavelength (km)	Vertical Wavelength (km) from the model of Vial and Forbes (1994)
January	59 (+/- 3)	41 (+/- 22)
February	31 (+/- 8)	14 (+/- 3)
March	16 (+/- 2)	14 (+/- 1)
April	24 (+/- 14)	14 (+/- 8)
May	40 (+/- 5)	-149 (+/- 112)
June	14 (+/- 2)	-121 (+/- 46)
July	16 (+/- 2)	-80 (+/- 66)
August	29 (+/- 9)	189 (+/- 63)
September	49 (+/- 14)	57 (+/- 4)
October	26 (+/- 4)	35 (+/- 2)
November	-21 (+/- 6)	240 (+/- 87)
December	53 (+/- 14)	121 (+/- 121)

Finally, note that the inferred vertical wavelengths are derived from the ~ 20 km height range, which is often only a small fraction of the wavelength. The results therefore should be regarded as representing a localised vertical wavelength, rather than implying a vertical wavelength existing over a large height range, or indeed an entire wavelength.

At times, when the amplitudes are small (particularly during the spring and early summer months), the phase estimates are less reliable. To clarify the behaviour of the phases during these times, seasonal-mean phases were calculated. The seasonal-mean phases are shown in Figure V.18, as are the seasonal means from the model. Comparing the observations to the model phases, it can be seen that agreement is excellent in autumn and winter (differences < 2 hours) and quite good in spring with a phase shift of ~ 7 hours, with the observations leading the model (which is also equivalent to a shift of 5 hours on the next phase track with the observations lagging behind the model). Significant differences occur in summer where the seasonal-mean observations have a much shorter vertical wavelength, than the near vertical modelled ones.

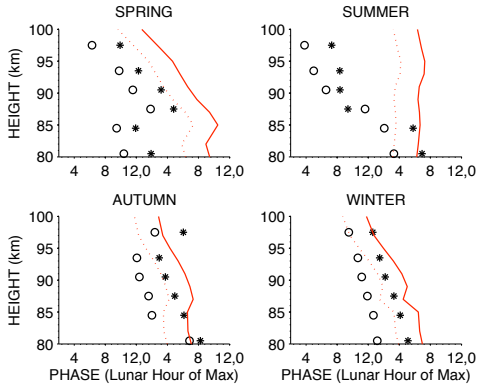


Figure V.18. Seasonal phase plots of the Lunar M_2 tide over Esrange. Black symbols indicate tidal phases calculated from the Esrange meteor radar over the period 1999 – 2007 and the red lines indicate results from the Vial and Forbes (1994) model. Solid lines show the zonal (stars) component and dashed show the meridional (circles).

V.4 Discussion

We have argued that a clear signature of the M_2 lunar tide is evident in meteor-radar observations of the mid-latitude and Arctic MLT region. In the introduction we noted that the periodogram of Figure V.2b actually displays *two* peaks at frequencies of ~ 1.9323 and 1.9350 cycles/day. The first of these is the M_2 lunar period. The second is possibly explained as a result of the response of the periodogram to a non-stationary time series (i.e., one in which the amplitude and phase of the M_2 oscillation vary over the course of the year). The seasonal cycle in M_2 tidal amplitude (large amplitudes in winter, small amplitudes in summer) could be interpreted by the periodogram analysis as the result of a beating between two closely-spaced frequencies. Since the amplitude varies with an annual cycle, the periodogram represents this as two oscillations, spaced in frequency such that one has one more cycle per year than the other. The difference in frequency between these oscillations is thus 1/365 cycles per day, which is ~ 0.00274 cycles/day and is the difference in frequency between the two peaks observed in Figure V.2b. We therefore assume that the two peaks in this figure, in fact, represent the response of the periodogram to the seasonally-varying M_2 lunar tide, and not two independent oscillations. This is one of the drawbacks of using this type of analysis. It would however, also occur in a Fourier based analysis. This is only a possible explanation of the dual peaks and may not necessarily be correct. If it were correct then the two peaks would most likely be centred over the lunar M_2 frequency. Instead, one peak is at the M_2 frequency and the other is at a slightly higher frequency.

A comparatively small number of studies have reported observations of the lunar semi-diurnal tide in the MLT region. The small amplitude of the tide makes it difficult to study and there appears to be a high degree of inter-annual variability.

Considering first the middle latitudes, Stening and Jacobi (2001) reported LF-D1 observations made over Collm (52°N, 15°E) during the 12-year interval 1985 - 1997. The location of Collm differs in longitude by 17° from the UK. This difference is accounted for using the analysis method described by Malin and Schlapp (1980) (i.e., the phases (the time of maximum) should be directly comparable for the migrating lunar tide). They describe their measurements as being representative of the height range ~ 90 – 100 km. The monthly-mean amplitudes agree well with those measured here by meteor radar. A similar seasonal behaviour is apparent with good agreement for the amplitudes. The phases again have a generally similar seasonal pattern, but the phases show a systematic difference of ~ 2 lunar hours, with the Collm data leading the UK data. This difference most likely reflects the slightly different height ranges observed by the two techniques.

Observations made by MF radar over Saskatoon (52°N, 107°W) in 1987 - 1988 were reported by Stening *et al*. (1994). The seasonal behaviour of the tidal amplitude was similar to that observed over the UK and reported here, however the amplitudes over Saskatoon were ~ 50% larger than over the UK. The Saskatoon results were obtained at a height of 99 km, higher than the representative meteor height of ~ 90 km considered here. There is also a known discrepancy with MF radar measurements over ~ 94 km where the measurements are influenced by group retardation effects, resulting in less reliable data (See Chapter VI and Manson *et al.*, 2004). So this again may explain the difference. We should also note that the high degree of inter-annual variability may lead to significant differences when only a few years of data are compared. Even considering this, the phases do generally agree to within ~ 2 lunar hours.

At Arctic latitudes, Stening *et al*. (1990) reported observations made over Poker Flat (65°N, 147°W) using an MST radar in meteor mode during the period 1983 to 1984. The monthly-mean amplitudes observed agree very well with those observed over Esrange, except during spring, when the Poker Flat amplitudes are large compared to those over Esrange (~ 3 - 5 ms^{-1} over Poker Flat c.f. ~ 1 ms^{-1} over Esrange). As before, the relatively short duration of observations at Poker Flat means that it is not clear if this is a systematic difference or a result of inter-annual variability.

Stening *et al*. (1990), also report tidal phases at heights of 90 and 103 km over Poker Flat. Comparing their phases for 90 km to those observed at the same height over Esrange, the phases are generally in good agreement from September - December, with the difference being less than ~ 2 lunar hours. However, significant differences exist in summer (June – August) and winter/spring (January – April), when the difference is ~ 6 lunar hours (i.e., the tides are effectively in anti-phase over the two sites).

Stening *et al*. (1990) also reported monthly-mean vertical wavelengths over Poker Flat for the months of March, July, September and December measured over the interval 1980 - 1984. Not all years produced useable data for each month. In March, one year of observations suggested a vertical wavelength of ~ 17 km (c.f. the Esrange six-year mean of ~ 16 $\pm$ 2 km). In July a five-year mean vertical wavelength was ~ 90 km (Esrange six-year mean ~ 16 $\pm$ 2 km), in September a two–year mean was ~ 51 km (Esrange six-year mean ~ 49 $\pm$ 14 km) and in December a one-year mean was ~ 29 km (Esrange six-year mean ~ 53 $\pm$ 14 km). The agreement between observations over Esrange and Poker Flat is thus excellent in March and September, but in July the vertical wavelength over Esrange is significantly shorter than over Poker Flat. In December, it is significantly longer.

Some of the interannual variability can be accounted for by the influence of the underlying stratospheric winds (Stening *et al*., 1997a). However, the small amplitude of the lunar tide in the MLT region effectively limits our study to monthly-mean properties and prevents the detailed investigation of the effects of stratospheric warmings etc. Nevertheless observations of the conjugate latitude in Antarctica may show different variability because of the lack of regular major sudden stratospheric warmings.

An interesting result is the strong similarity of the seasonal behaviour of the M_2 lunar tide and the solar 12-hour S_2 seen in figures V.14 and 15. More detailed observations of the Solar S_2 tide over Esrange have been reported by Mitchell *et al*. (2002). Both tides have strong amplitude maxima in winter and brief secondary maxima in September. Such similarity may simply reflect the very similar propagation environments encountered by the lunar and solar tides which are very close in frequency and global-scale structure. Possible causes of this seasonal variability in the S_2 solar tide have been discussed by Riggin *et al*. (2003) – including refractive effects caused by shears in the mean zonal wind and gradients in temperature. Given the similarity between the solar and lunar tide, it is likely that propagation conditions, rather than source variations, are the mechanisms that produce the general variability observed in the tides. If the propagation conditions are the source of the variability,

then they are likely to affect the solar and lunar migrating tides in similar ways, but if source changes (e.g., concentrations of ozone or water vapour) are the cause of the solar S_2 tidal variation, then seasonal variations are likely to differ between the two tides.

It should also be noted that seasonal variation of the semi-diurnal tide at high latitude is likely to be partly due to the combination of non-migrating modes. An example of non-migrating solar semi-diurnal tidal components in the southern hemisphere at high latitudes is given by Murphy *et al*. (2003).

An important consequence of the significant amplitudes reached by the lunar tide in the MLT region, which have often been ignored in analyses, may be in the interpretation of tidal variability. Quasi-periodic variations in the amplitude of tides are often taken to result from non-linear interactions between tides and planetary waves. In particular, modulation of tidal amplitude at the period of a planetary wave may result from non-linear mixing generating secondary waves, which then beat with the tide (Teitelbaum and Vial, 1991). Examples include interaction between the 12-hour S_2 solar tide and the quasi-2-day planetary wave (Mitchell *et al*., 1996; Palo *et al*., 1999). Here we should note that a simple beating between the 12.42- hour M_2 lunar tide and the 12 hour S_2 solar tide will result in an apparent modulation of the S_2 tide at a period defined by the difference in frequency between the M_2 and S_2 tides (a frequency of ~ 0.0677 cycles per day, equivalent to a period of ~ 14.8 days). This latter period is close to that of the 16-day planetary wave and so may be mistakenly interpreted as resulting from a non-linear interaction between the 16-day planetary wave and the 12-hour S_2 solar tide (a point noted by Stening *et al*., 1997b; Pancheva *et al*., 2002b). Care must therefore be taken in attributing apparent fluctuations of the 12-hour S_2 solar tide at a period near the 16 day wave to interaction with the 16-day wave.

We may therefore ask the question: “Does the lunar tide detected in the meteor radar data presented here actually arise from an interaction between the 16-day wave and the solar 12-hour tide?”. We are confident that we have identified a genuine lunar tidal signal for the following two reasons. Firstly, the observed monthly-mean lunar tidal phases are consistent from year to year. This would not be so if the apparent lunar tide was actually being produced by interaction between the 12-hour solar tide and 16-day planetary wave. This is because the planetary wave is randomly phased and so would produce a randomly phased apparent lunar tide. Secondly, the 16-day wave actually displays a range of periods near 16 days (e.g., Mitchell *et al*., 1999; Luo *et al*., 2002) and so any generation of an apparent lunar tide would actually result in a rather broad spectral peak near lunar tidal frequencies. However the periodogram of figure V.2a

reveals a very sharp peak at lunar tidal frequencies. We therefore conclude that the analysis presented here has detected a genuine lunar tidal signal in the data.

V.5 Conclusions

Two meteor radars have been used to measure the lunar M_2 tide at mid (UK, 52°N) and Arctic latitudes (Esrange, 68°N). The observations over the UK comprise the longest study of the lunar tide in the MLT region to date, covering an interval of 16 years. The observations of the lunar tide at Esrange make up the longest continuous data set at Arctic latitudes and are currently the most northerly study of the lunar tide.

At both locations the lunar tidal amplitudes were found to be largest in winter (up to ~ 11 ms^{-1}), with a secondary maximum in autumn. Data from Esrange shows the amplitudes increase with height between 78 and 98 km. Generally larger amplitudes were found over Esrange than over the UK, which is surprising due to the latitudinal structure expected from the Hough modes, which can be seen in the model. Annual-mean amplitudes from the UK data set show that the M_2 lunar tide exhibits significant year-to-year variability.

The zonal and meridional phases observed over both locations are in quadrature, with the meridional component leading the zonal component. As both components have similar amplitudes this indicates that the tide is circularly polarised and rotating clockwise viewed from above. For the northern hemisphere this is in agreement with the Coriolis force. The phases seen over Esrange also decrease with increasing height, producing a spiralling wind vector which rotates clockwise and downward in height. Downward phase propagation corresponds to an upward energy flow, from classical tidal theory (Chapter II). Therefore tides with a downward clockwise (anti-cyclonic) spiralling wind vector, as in the Esrange case presented here, are likely to originate from sources below.

Over the UK, results from the Vial and Forbes, 1994 model resolved the seasonal structure very well, but predicted the amplitudes to be approximately a factor of 2 larger than those observed. This is most likely due to the rapid increase in amplitude with height. The seasonal pattern in phases also fits very well with that observed, but there is a systematic offset of ~ 2 lunar hours.

The results from Collm (Stening and Jacobi, 2001) show a very good agreement with the amplitudes, but an approximate 2 lunar hour phase difference, most likely due to

uncertainty in height using the LF-D1 method. The data from Saskatoon (Stening *et al*., 1994) show the same seasonal behaviour, however, the amplitudes were ~ 50% larger. This difference in amplitude can be explained by the difference in height of the measurements. The Saskatoon observations were made at a higher altitude than those of the UK. Comparisons of the phases between these sites show very good agreement.

For Arctic latitudes, the Vial and Forbes (1994) model amplitudes agree very well with the observations presented here. However, the observations reveal smaller amplitudes than predicted by the model in autumn and early winter. The model reproduces the phase behaviour very well.

Vertical wavelengths measured in the Arctic range from 16 km in March to 59 km in January. This disagrees with the model's vertical wavelengths during summer, where much longer wavelengths are predicted. Note that this is where the errors on the phases are at their highest.

Comparisons with observations from Poker Flat (Stening *et al*., 1990) show reasonable agreement in the amplitudes to those observed over Esrange. The phases agree during September to December, but, in the majority of other months, phases are in anti-phase to those over Esrange, which could reflect the presence of non-migrating lunar tidal modes.

A long term average of Poker Flat data (1980 – 1984) shows a long vertical wavelength in July, but in a single year's data, the vertical wavelength is shown to be much shorter and more in agreement with that seen over Esrange. This may be due to the fact that the tide exhibits a high degree of interannual variability.

The seasonal variation of amplitude and phase seen in the tide over both locations has been found to be very similar to that seen in the S_2 solar semi-diurnal tide, including the very abrupt autumnal equinox enhancement in amplitude.

As a final note of caution, a beating between the solar S_2 and the lunar M_2 tides could produce an apparent modulation of the S_2 tide at periods near 16 days. This could be misinterpreted as a non-linear interaction between the 16-day planetary wave and the solar S_2 tide.

Note that the work presented in this Chapter was published as: Sandford *et al*., 2006.

Chapter VI

VI. Migrating and Non-migrating Lunar Tide in the Equatorial and Antarctic MLT Regions

Non-migrating tides, like migrating tides are global-scale oscillations with periods that are harmonics of a solar or lunar day. However, they do not follow the apparent motion of the forcing body, (i.e., the Sun or Moon); they are not Sun-synchronous. Unlike migrating tides, they may be stationary, propagate eastward, or propagate westward. The only constraint on their characteristic zonal wavenumber is that they not equal that of the migrating wave. The first attempts to model non-migrating tides were those of the solar tide, which assumed that their source was the nonuniform global distributions of water vapour and zonally asymmetric components of near-surface heating (Kato, 1989; Hagan et al., 1997). Other sources of solar non-migrating tides are thought to include zonal variations in surface properties and wave-wave nonlinear interactions, whose relative contributions remain unexplored (Moudden and Forbes, 2008).

Observations of non-migrating components of the semi-diurnal solar tide began in Antarctica with Hernandez *et al*. (1993), when a weak zonal wavenumber 1 mode was identified at the South Pole using hydroxyl airglow measurements. The study was then extended by Forbes *et al*. (1995) and Portnyagin *et al*. (1998) who used meteor radar wind measurements, which also covered the South Pole area. Portnyagin *et al*. (1998) suggests that non-migrating components of the semi-diurnal tide are able to extend northward to at least 68°S. Non-migrating diurnal tides have also been measured, including equatorial observations (Lieberman, 1991 and Oberheide and Gusev, 2002).

As for the lunar tide, the tidal potential is very well determined, however, this potential also affects the Earth and oceans. This modifies their mass distributions and elevations creating secondary tidal potentials (Vial and Forbes, 1994). The first evident effect of oceanic tides on atmospheric tides is the generation of non-migrating modes. As the oceanic tides are constrained by the presence of continental coasts, the ocean tides cannot be fully described by zonal wavenumber two alone. The load tides also

have a significant effect in producing non-migrating modes. The load tide is induced by the distortion of the Earth's surface from the overlying ocean (i.e. when the ocean tide is positive; it creates a load on the underlying crust which is depressed). This load tide is generally out of phase with the ocean and atmospheric lunar tides.

In the past, the limited spatial coverage of radars, has lead to the assumption of longitudinal independence of the tides. This is appropriate in the case of migrating tides, but does not take into account that there may be non-migrating components. Satellites do not have this limitation, as they generally have large spatial coverage but, they do however have poor local time coverage, which limits the study of tides to very long epochs. Observations from a number of different longitudinal stations which have similar latitudes are therefore very important for studying non-migrating tides. Radars for example contain the necessary temporal resolution to resolve the tides and the use of multiple stations is a good start in identifying possible non-migrating tides.

VI.1 Equatorial Studies

A number of radar stations have measured the lunar M_2 tide to reveal the seasonal behaviour. With the addition of measurements from the Ascension Island (8°S, 14.4°W), a longitudinal chain of radars around the equatorial region is available. These radars are almost equally spaced in longitude, thereby allowing the consideration of non-migrating tidal components by virtue of this longitudinal separation. The Ascension Island (8°S, 14.4°W), Christmas Island (2°N, 157°W) and Jakarta (6°S, 253°W) radar systems all lie within 10° of latitude over the equator. Christmas Island and Jakarta are the closest together being space by ~96° in longitude. Ascension Island is approximately halfway between the two on the opposite side of the Earth, as shown in Figure VI.1.

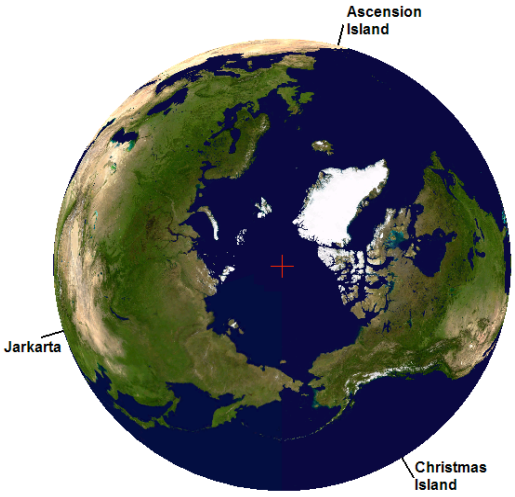


Figure VI.1: View of the Earth from the North Pole, the location of the three equatorial stations marked, indicating the longitudinal separation of the stations.

Measurements of the equatorial lunar M_2 tide have been observed over the two sites of Christmas Island (2°N, 157°W) and Jakarta (6°S, 253°W) by Stening *et al*., (1997b and 2003, respectively). Unfortunately the radars do not contain common periods of observation, however the general seasonal cycle has been attained from a number of years of data. A new set of observations here have been collected by a meteor radar on Ascension Island (8.0ºS, 14.4ºW) in the equatorial Atlantic region. The radar is a SKiYMET system, similar to that described in Chapter III. It operates at a frequency of 43.5 MHz. Peak power is either 12 or 6 kW (post October 2005). Here we consider data recorded from May 2001 to April 2007. Horizontal winds at heights of 78 – 100 km are calculated in the independent height gates with centres at 81.1, 84.6, 87.5, 90.4, 93.3 and 96.8 km.

Monthly-mean tidal amplitudes and phases have been calculated for each height gate using the least squares method of Malin and Schlapp (1980), the same as that used in Chapter V. Because of the high spectral resolution required to separate the lunar M_2 and solar S_2 tides, the analysis was again conducted on a monthly-mean basis. Only results from months with more than 16 days of data were accepted. This criterion means that 22 out of 76 months of data were discarded.

VI.1.1 Ascension Island Results

Monthly-mean tidal amplitude and phase were calculated in each height gate for all of the available months of the 6-year dataset. The monthly-means for each month from successive years were then vector averaged to yield a composite year similar to the observations of Stening *et al*. (1997b and 2003). Figure VI.2a, b presents the resulting

zonal and meridional a) amplitudes and b) phases, averaged over all six height gates, i.e., 78 – 100 km. These figures are directly comparable with Figures V.11a, b for the UK and Figures V.13a, b for the Arctic.

The monthly-mean amplitudes of Figure VI.2a are seen to be mostly between ~ 1.5 and 3 ms^{-1}. During most of the year the meridional amplitudes are larger, but around the equinoxes the zonal amplitudes are larger. The meridional amplitudes exhibit two distinct equinoctial minima. The phases of Figure VI.2b reveal that the zonal component leads the meridional throughout the year, i.e., an anticlockwise rotation with time. The phase difference between the two components is ~ 3 hours in Southern-Hemisphere (SH) summer (November to March), which changes at the equinoxes to values as large as six hours for the winter months. To more easily visualise the relationship of amplitude, phase and polarisation, Figure VI.2c presents hodographs of the vector averaged monthly-mean properties. As can be seen from the figure, SH summer months are more circularly polarised. The transition regions (March and October) are elliptically polarised towards the zonal and for the majority of the rest of the year (May – September), the tide is elliptically polarised in anti-phase.

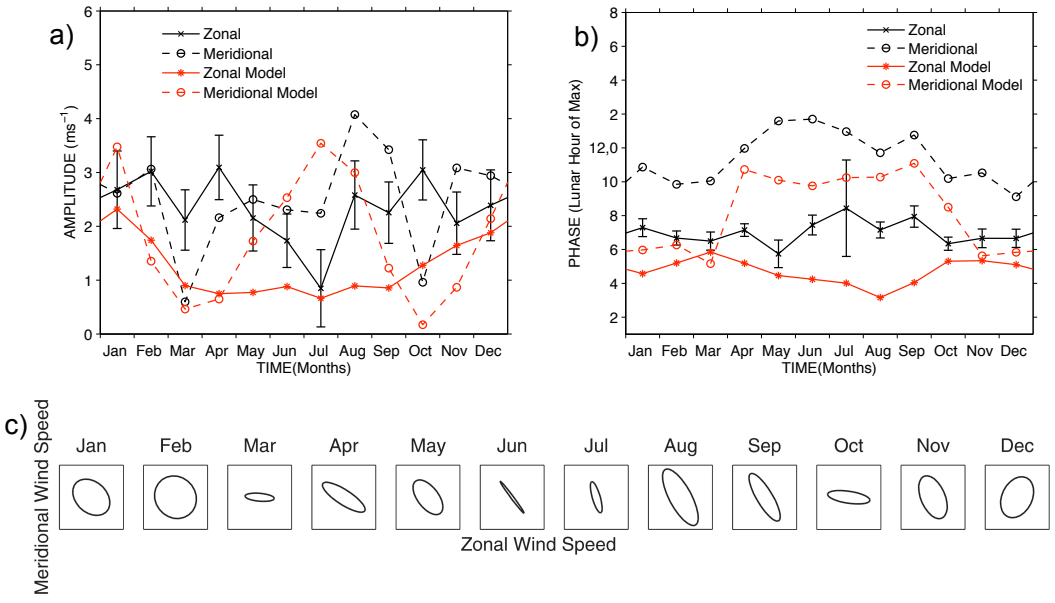


Figure VI.2a, b and c. Monthly-mean a) amplitudes, b) phases and c) hodographs, of the lunar M_2 tide calculated from Ascension Island data (2001 – 2007) vector averaged to produce the seasonal variation indicated in black. Red lines indicate amplitude results from Vial and Forbes (1994) model between 80 and 100 km. Solid lines represent the zonal component and dashed the meridional. Error bars indicate the standard error on the mean.

To investigate the vertical structure of the tide, the monthly-mean amplitudes in each height gate are presented as contours in Figure VI.3a, b for the zonal and meridional components respectively. The contours reveal that in most months the tidal amplitudes

increase with height to values as large as ~ 10 ms^{-1} at heights of ~ 97 km. As suggested by the all-height-gate means of Figure VI.2, the wintertime and summertime meridional amplitudes are generally larger than the zonal ones. Around the equinoxes, the amplitudes fall to very small values, often < 2 ms^{-1} at all heights. Interestingly, in February and near the autumnal equinox the zonal amplitudes display a short-lived maximum that peaks at a height of ~ 88 – 90 km. This feature is reminiscent of the local autumnal maxima of the 12-hour solar tide, sometimes observed at this and other latitudes.

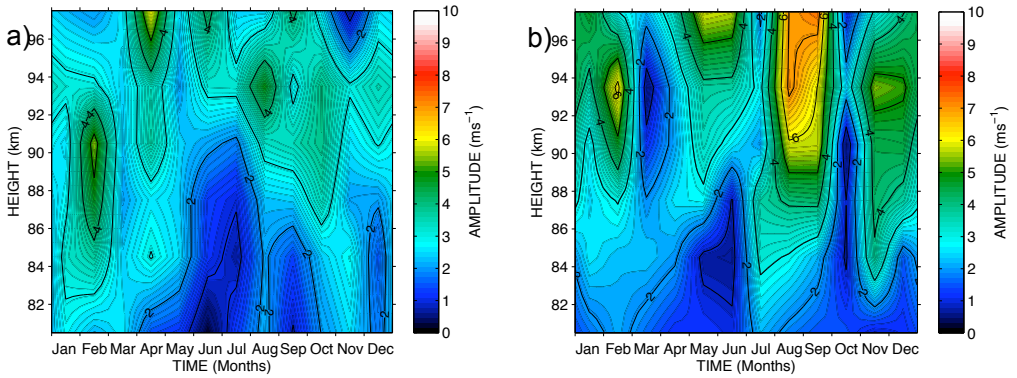


Figure VI.3a, b. Contour plot of the lunar M_2 tidal amplitude calculated from monthly-mean amplitudes (vector averaged to give the seasonal variation) of a) zonal and b) meridional wind data over the period 2001 – 2007 from the Ascension Island meteor radar. Colour scale indicates amplitude.

To compare this behaviour with that of the solar S_2 semi-diurnal tide, the same analysis was applied. Figures VI.4a and b present amplitude contours of the a) zonal and b) meridional components of the solar S_2 tide. The data have been vector averaged to produce a composite year from the six years of available data (May 2001 – April 2007). Both of the figures of Figure VI.4 reveal very different behaviour to that seen in Figure VI.3. There is the similarity that the zonal component has smaller amplitudes than the meridional, as seen with the lunar M_2 tide, and they also tend to increase in amplitude with height. However, the solar tide tends to maximise in March-April in both components, with a secondary maximum in October (much more prominent in the meridional component). The lunar tide has localised maxima in February and August. This also provides further evidence that the lunar tidal signal is not just a sideband of the solar S_2 tide, from their very different behaviour.

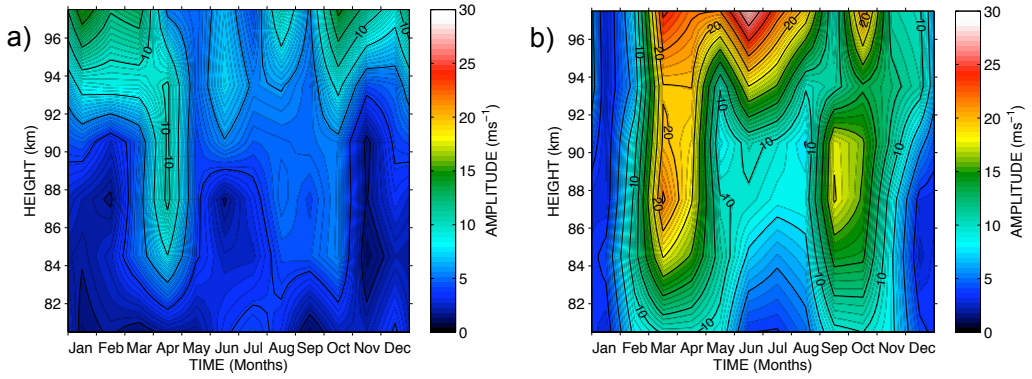


Figure VI.4a, b. Contour plot of the Solar S_2 tidal amplitude calculated from monthly-mean amplitudes (vector averaged to give the seasonal variation) of a) zonal and b) meridional wind data over the period 2001 – 2007 from the Ascension Island meteor radar. Colour scale indicates amplitude.

Figure VI.5 presents height profiles of zonal and meridional amplitude for each month of the Ascension Island observations along with the corresponding predictions of the Vial and Forbes (1994) model. In both components, the observed amplitudes are either approximately similar to or significantly larger than the model predictions. The differences are largest in autumn (February to April) and spring (August to September) where the observed amplitudes rise toward values near 10ms^{-1} at the upper heights, but where the model predictions are usually somewhat smaller.

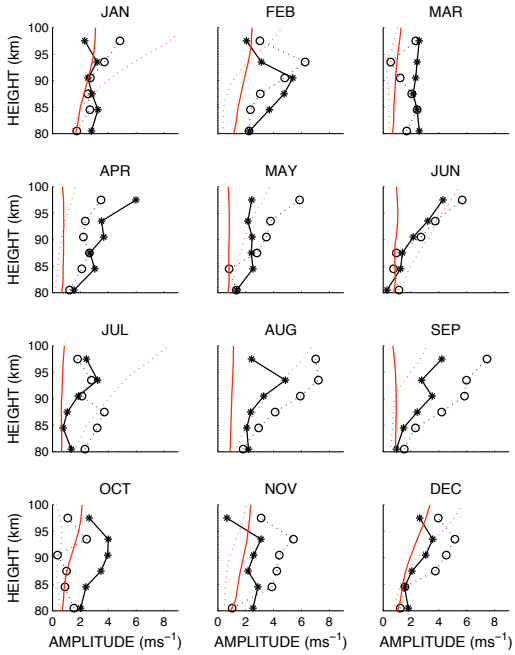


Figure VI.5. Monthly-mean amplitudes of the lunar M_2 tide. Black lines/symbols indicate tidal amplitudes calculated from the Ascension Island meteor radar over the period 2001 – 2007 and the red lines indicate results from the Vial and Forbes (1994) model. Solid lines show the zonal (stars) and dashed the meridional component (circles).

Figure VI.6 presents a similar analysis of the phases. The observed phases generally decrease with height, corresponding to an upwardly-propagating tide. The zonal phase leads the meridional by amounts varying from < 1 lunar hour in March to ~ 10 lunar hours in May. The observed zonal phases lag the model phases at all heights and in all months. The differences are largest around the summer solstice (January) and in winter and early spring (June to September) when the differences can exceed 6 hours. The differences are rather less in the case of the meridional phases. Uncertainties in the measured phases are not shown for reasons of clarity. However, the mean uncertainty in phase over the entire ensemble of monthly data is 1.3 hours and the largest uncertainties occur when the amplitudes are small.

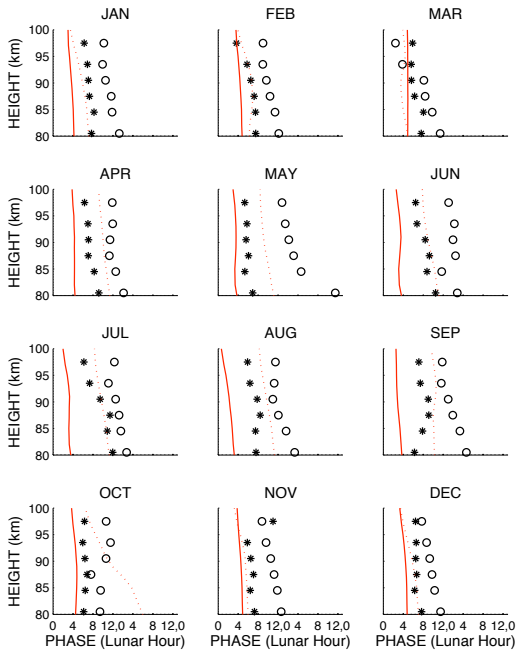


Figure VI.6. Monthly-mean phases of the lunar M_2 tide over Ascension Island. Black symbols indicate tidal amplitudes calculated from the Ascension Island meteor radar over the period 2001 – 2007 and the red lines indicate results from the Vial and Forbes (1994) model. Solid lines show the zonal (stars) component and dashed show the meridional (circles).

The variation of tidal phase with height allows calculation of the vertical wavelength for each month. Table VI.1 lists the observed vertical wavelengths and those from the model of Vial and Forbes (1994). The vertical wavelengths are an average of the zonal and meridional components. In the zonal and meridional component individually, the vertical wavelengths are generally between ~ 30 and 80 km but range from a minimum of 18 ± 8 km in November to a maximum of 1509 ± 516 in October, both in the zonal component. The model predicts meridional wavelengths roughly comparable to these values, but predicts significantly larger zonal vertical wavelengths (> 150 km). The

observations over Ascension Island and the model both exhibit vertical wavelengths that can differ between the two components.

Table VI.1. Monthly-mean vertical wavelengths observed over Ascension Island (8°S, 14°W) from May 2001 to April 2007 and those from the model of Vial and Forbes (1994). The wavelengths are calculated over the height range 80 – 100 km and zonal and meridional values averaged. Errors indicate the error on the fit to the gradient.

Month	Measured Vertical Wavelength (km)	Vertical Wavelength (km) from the model of Vial and Forbes (1994)
January	79 (+/- 7)	97 (+/- 11)
February	55 (+/- 4)	104 (+/- 41)
March	33 (+/- 4)	178 (+/- 123)
April	61 (+/- 20)	141 (+/- 19)
May	28 (+/- 8)	125 (+/- 15)
June	50 (+/- 23)	107 (+/- 7)
July	41 (+/- 6)	106 (+/- 7)
August	66 (+/- 7)	84 (+/- 4)
September	64 (+/- 18)	296 (+/- 60)
October	-83 (+/- 50)	37 (+/- 2)
November	30 (+/- 8)	123 (+/- 24)
December	84 (+/- 13)	93 (+/- 12)

VI.1.2 Discussion of the Equatorial Results

Only two studies of the lunar M_2 tide in the low-latitude MLT region have been reported. Firstly, Stening *et al*. (1997b) presented MF-radar measurements made over Christmas Island (2°N, 203°E) in the Pacific sector from 1990 - 1993. These results along with the Ascension Island and Jakarta results are shown in Figure VI.7 for Amplitude and in Figure VI.8 for phase, a) for the zonal component and b) for the meridional. The model outputs of the Vial and Forbes (1994) Model at 8°S are also indicated on these figures.

The amplitudes over Christmas Island were broadly similar to those reported here. However, Stening *et al*. observed a systematic, year-long, 6-hour phase difference between the zonal and meridional components and saw no equinoctial transition, unlike the seasonally-varying meridional phases reported here. Measured in lunar hours, there is also a systematic phase difference of ~ 3 hours in the zonal components, with Ascension Island lagging Christmas Island.

Secondly, Stening *et al*. (2003) reported observations made by meteor radar over Jakarta (6.4°S, 107ºE) between 1987 and 1988 (also included in Figures VI.7 and VI.8). The monthly-mean amplitudes at Jakarta are similar to those over Ascension Island during summer, but at other times the latter amplitudes are larger, particularly

during late winter (August and September) in the meridional component. Stening *et al*. (2003) also noted that the phase relationship between the zonal and meridional components of the tide over Jakarta in southern spring and summer was more characteristic of a northern hemisphere lunar tide – a type of behaviour not seen over the slightly more southerly Ascension Island (i.e., the meridional component leads the zonal, giving a clockwise rotation). Comparing Jakarta to Ascension Island in summer and autumn, the zonal phases agree to within ~ 2 hours. However, in SH winter and early spring (June to September) the phases over Ascension Island do not change significantly, whereas, at Jakarta the zonal phases change by ~ 6 hours. During this time the zonal phases over the two sites are thus in approximate anti-phase. In contrast, the meridional phases have a similar seasonal progression with a systematic difference of ~ 4 hours (Jakarta leading Ascension Island, with the exception of April).

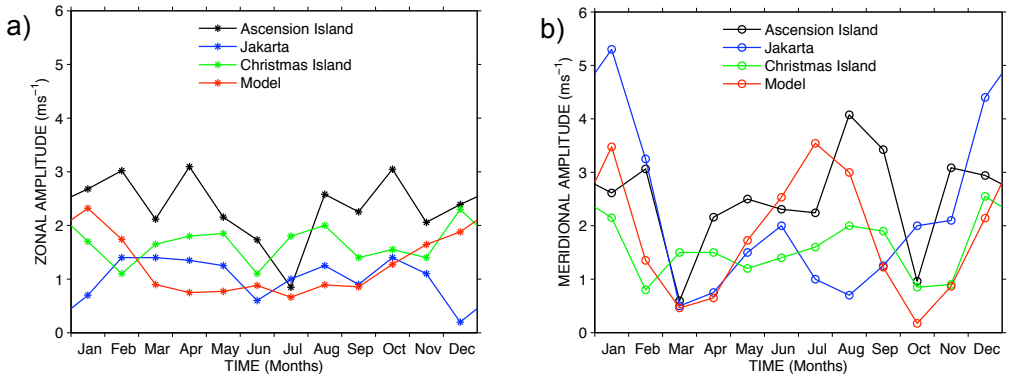


Figure VI.7a, b: A yearly average of monthly-mean lunar M_2 tidal amplitude over Ascension Island (black), Jakarta (blue) and Christmas Island (green). Amplitudes are calculated using different time spans for each site. The red lines indicate results from the Vial and Forbes (1994) model. Zonal (stars) component in figure a) and meridional (circles) component in b). (Using data from Stening *et al*., 1997b and Stening *et al*., 2003).

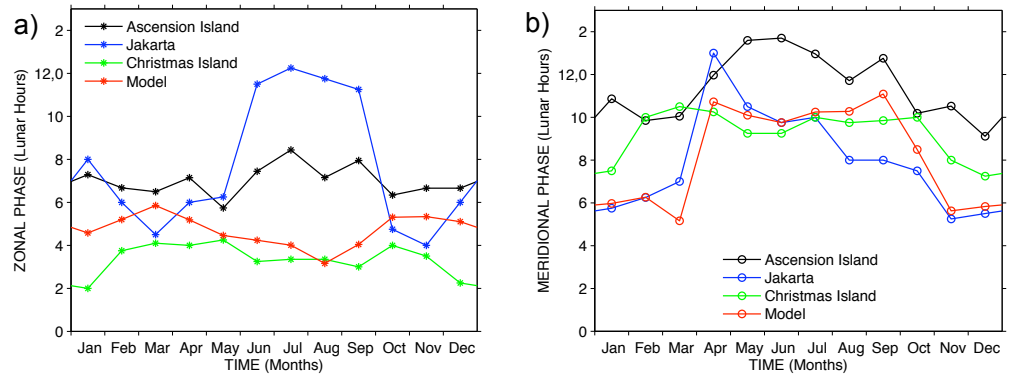


Figure VI.8a, b: A yearly average of monthly-mean lunar M_2 tidal phase over Ascension Island (black), Jakarta (blue) and Christmas Island (green). Phases are calculated using different time spans for each site. The red lines indicate results from the Vial and Forbes (1994) model. Zonal (stars) component in figure a) and meridional (circles) component in b). (Using data from Stening *et al*., 1997b and Stening *et al*., 2003).

Stening *et al*. (2003) also examined the latitudinal behaviour of the equatorial lunar M_2 tide in a seasonal version of the Global Scale Wave Model (GSWM). The model suggests zonal phases vary little with latitude throughout the year within 10º of the equator. All three sites considered here should thus see similar zonal phase behaviour – in contrast to the results described above. Meridional phases also vary little with latitude, except in NH winter (December-February) when they change by approximately 4 lunar hours between 10°N and 10°S. The zonal phases observed at 90 km over Ascension Island are actually very close to these predicted GSWM phases, presented by Stening *et al*., 2003 (with less than one hour difference in 10 months of the year). The meridional phases only agree well between April and September.

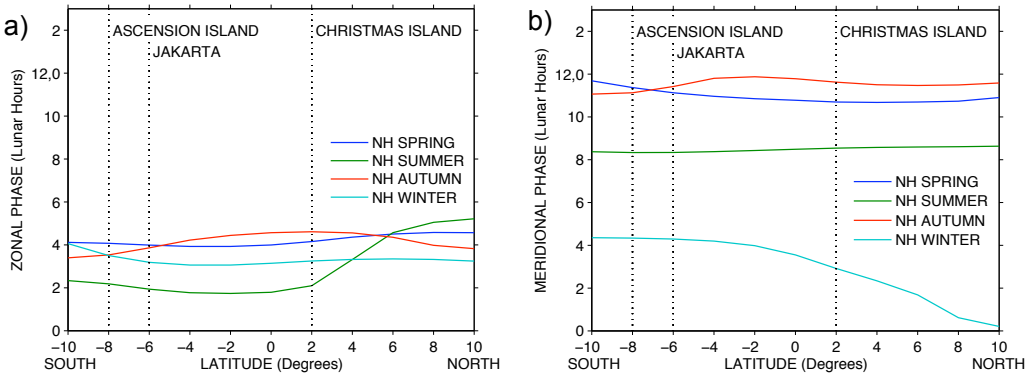


Figure VI.9a, b: Phase of the Vial and Forbes (1994) lunar M_2 tidal Model for 97 km between 10°S and 10°N (After Stening *et al*., 2003).

Possible explanations for the large differences between Ascension Island and the other two sites include:

1) Inter-annual variability, because the various studies are not contemporaneous.

2) Biases between different radars/techniques which may make direct inter-comparison of observations difficult (e.g., Manson *et al*., 2004). However, such biases are more likely to affect comparisons of tidal amplitude rather than the phases considered here.

3) Non-migrating lunar M_2 tides may be present, resulting in significant longitudinal variations of amplitude and phase. Vial and Forbes (1994) suggests that such non-migrating modes are significant, although they do not present results for low latitudes in the MLT. The systematic nature of some of the observed phase differences is in agreement with what might be expected in the presence of non-migrating modes.

VI.1.3 Conclusions of the Equatorial Studies

The lunar M_2 tide in the MLT over Ascension Island (8.0°S, 14.4°W) can reach amplitudes of ~ 8 ms^{-1} and the amplitudes often increase with height. A clear seasonal

behaviour exists with maximum amplitudes in SH winter and late summer, with much smaller amplitudes in the zonal component than the meridional. Agreement is good with the Vial and Forbes (1994) model, except that larger amplitudes are observed than predicted (up to a factor of two) and there is a systematic phase offset of ~ 2 hours with the observations lagging the model.

Comparisons with observations made in earlier years over Christmas Island and Jakarta show a reasonable agreement in the case of tidal amplitudes but the phase exhibits very significant differences - sometimes becoming so large they result in an anti-phase relationship between Ascension Island and Jakarta. Non-migrating lunar tides may account for some of these differences.

VI.2 Antarctic Studies

Of the various lunar tidal modes, only the lunar semi-diurnal M_2 mode, which has the period of ~ 12.4206 hours, appears to reach significant amplitudes in the Antarctic MLT region. To date there has been only one study of the lunar tide at high southern latitudes. This was the study of Stening *et al*. (1995) who investigated the M_2 tide over Scott Base (78°S) using MF-radar observations made over the interval 1983-1984. This study suggested that the M_2 tide has largest amplitudes in the winter months (of up to ~ 8ms^{-1}).

In the work presented here, data from two radars in the Antarctic are used to investigate the M_2 lunar tide at high southern latitudes. These two radar stations have coincident latitude and also the measurements are contemporaneous. This is therefore ideal for studying the tides in more depth. One radar is an all-sky meteor system and is located at Rothera (68°S, 68°W). This radar has been operating since February 2005. Data recorded over the interval February 2005 to August 2007 is considered. The other system is a MF radar located at Davis (68°S, 78°E). This radar has been operating since May 1994. Data recorded over the interval May 1994 to September 2006 is considered. The lunar M_2 tidal model of Vial and Forbes (1994) will also be examined.

The University of Bath/British Antarctic Survey - Rothera meteor radar is similar to that described in Chapter III. The radar has an identical specification to that of the Esrange meteor radar and is similar to that of the Ascension Island meteor radar; the only difference being the transmission frequency, which affects the ceiling height of meteor detections as described in Chapter III. The system operates at a transmission

frequency of 32.5 MHz and uses the all-sky configuration. Horizontal winds are calculated in six independent height gates which again are centred at heights of 81.1, 84.6, 87.5, 90.4, 93.3 and 96.8 km.

The Davis radar is owned and operated by the University of Adelaide and the Australian Antarctic Division (AAD) in conjunction with the University of Adelaide and Atmospheric Radar Systems Pty. Ltd.. It is a MF (Medium Frequency) radar system operating in the spaced-antenna configuration (Vincent, 1984). It uses the method of spaced antenna drifts to measure the winds in the MLT region. This radar system has been in continuous operation since May 1994 and measures the horizontal wind speed and direction at heights of ~ 70 to 98 km. The routine height resolution is 2 km. The data obtained above 94 km from the MF radar may be affected by reflections from the D-region of the ionosphere, so analysis of the MF-radar data here will only consider heights up to 92 km (Manson *et al*. 2004).

To confirm that a lunar tidal signal can be identified in the Antarctic data, Figure VI.10a, b presents a Lomb-Scargle analysis similar to that of Figure V.2a. Figure VI.10a includes data from Rothera and VI.10b from Davis. In each case the analysis is based on the twenty months of data, 2005-06 for Rothera and 1997-98 for Davis. In the case of Rothera, the figure presents results from 95-100 km (the 96.8 km height gate) and in the case of Davis the results are from 89-91 km. The figure shows the presence of strong signals of period 24, 12 and 8 hours, which are the familiar diurnal S_1, semi-diurnal S_2 and ter-diurnal S_3 solar tides. Also evident in each figure is a significant but small signal at the expected 12.42-hour period of the lunar semi-diurnal M_2 tide. However, this analysis contains no information about the phase of its signal or how the tide varies with height, so a superposed-epoch analysis was also carried out.

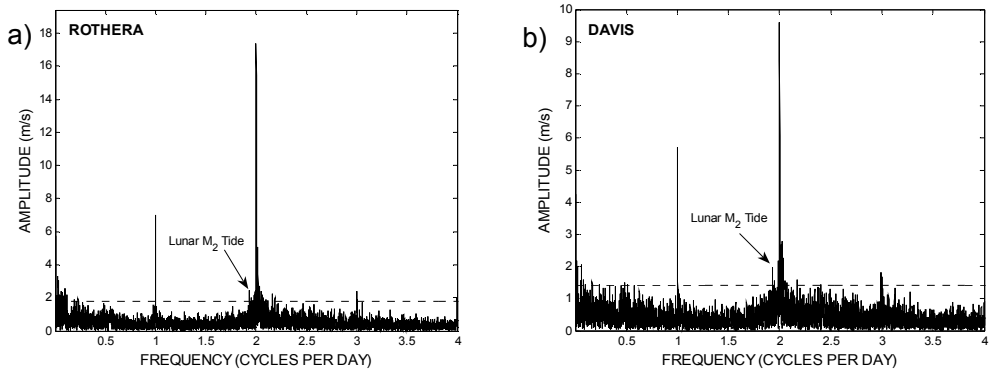


Figure VI.10 a, b: Lomb-Scargle analysis of meridional winds for a) Rothera at a height of 97 km over the interval February 2005 to December 2006 and b) Davis at a height of 90 km over the interval January 1997 to December 1998. Both plots show the 95% significance level indicated by the dashed line.

Figure VI.11a, b presents a long term superposed-epoch analysis of the lunar day for all 13 years data from Davis. Using this analysis, any coherent feature which repeats itself with a period equal to exactly n times the lunar transit frequency (where n is an integer) will be evident, whereas any oscillations present with frequencies not related to the lunar-transit frequency will self cancel and therefore not be evident. Figure VI.11a presents this analysis applied to the zonal component and Figure VI.11b presents it for the meridional component. Note that at each height gate the mean has been subtracted from the time series so that any oscillations present are not lost against large changes with height resulting from the mean flow. The plots are shown in lunar time (a lunar day being the time it takes for the Moon to make one complete orbit of the Earth). In each figure a clear wave-like structure is evident across the entire height range of 70 – 90 km considered. Two cycles are evident per lunar day, suggesting that this feature is the lunar M_2 tide. The amplitude of the feature seems to be ~ 0.8 ms^{-1}. Further, a clear phase progression is observed, decreasing with height. This, in the case of a wave or tidal motion, corresponds to an upwardly-propagating wave or tide as described in Chapter II. Note that because this analysis uses the entire 13-year data set from Davis, seasonal fluctuations in the phase of the wave-like feature will lead to a degree of self-cancellation, so the wave amplitudes revealed are very much lower limits.

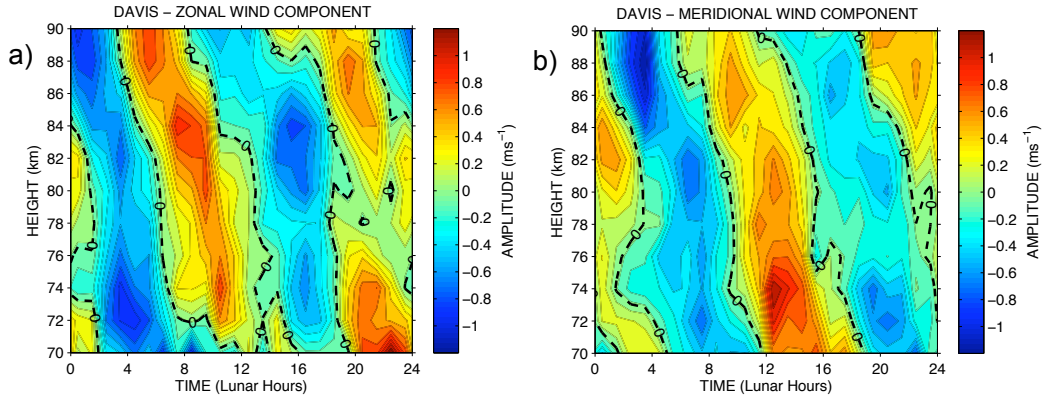


Figure VI.11a, b: Composite lunar day of the winds over Davis for the interval 1994 to 2006, measured from the time of the first new Moon in this interval for a) zonal winds and b) meridional winds. The zero line is marked by the dashed line.

Similar to Figure V.5, to check for spectral leakage from the large-amplitude solar semi-diurnal S_2 tide to the lunar M_2 tidal frequency, an analysis similar to that above, but using a superposed epoch period of 23.16 hours instead of 24.84 hours was carried out. Figures VI.12a, b show the results of this analysis, presented in a similar format to that of Figures VI.11a, b. If it were the case that the wave-like features of Figure VI.11a, b were caused by spectral leakage from the solar S_2 tide, then we might expect a similar leakage into this higher-frequency analysis. However, although there

is some variability within the composite day, it does not have the clear wave-like features of Figure VI.11 and those features that are present are of smaller amplitude (~ 0.4 ms-1, c.f. ~ 0.8 ms^{-1}).

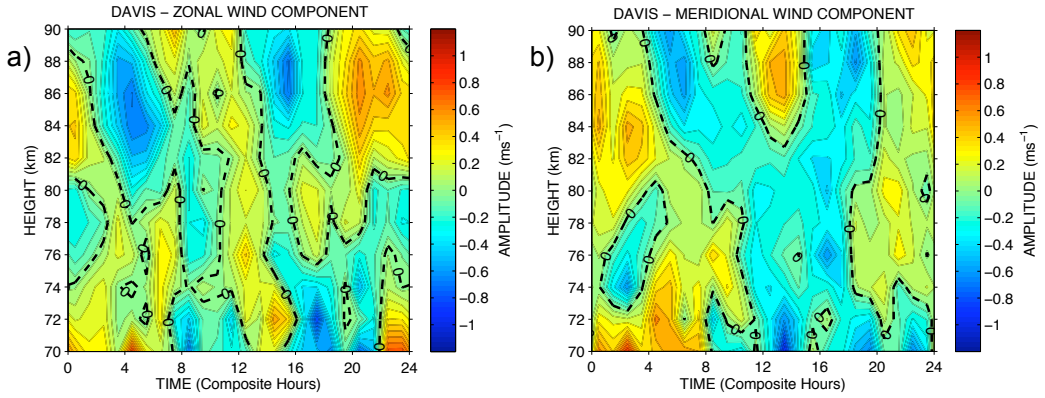


Figure VI.12a, b: As for Figure VI.11, but using a composite day of 23.16 (solar) hour duration.

The spectral and composite lunar-day analysis reveals a coherent wave-like feature with two cycles per lunar day. The feature exhibits a phase gradient corresponding to that expected for an upwardly-propagating wave or tide and based on the analysis of searching for similar features spaced at an equal frequency from the solar S_2 tidal frequency, suggests this feature is not caused by spectral leakage from the solar tide. It should also be borne in mind that non-linear interactions between the S_2 solar tide and the "16-day" planetary wave can generate secondary waves with frequencies very close to that of the lunar M_2 tide – in particular the "difference" secondary wave resulting from a 12-hour tide / 16-day wave interaction will have a period near 12.39 hours, c.f. the M_2 period of 12.4206 hours. Such secondary waves have been proposed to account for some observed aspects of tidal variability (e.g., Teitelbaum and Vial, 1991; Pancheva *et al*., 2002a; Luo *et al*., 2002). However, there is no *a priori* reason why such secondary waves would have any phase relationship with the position of the Moon (unlike lunar tides which by definition will). Over the course of several years, the random phasing of the secondary waves with respect to the position of the Moon will lead to self-cancellation and so such waves should not produce a significant signal in the superposed-epoch analysis of Figure VI.11. We therefore conclude that the wavelike features evident in the superposed-epoch analysis of Figure VI.11 and the spectral analysis of Figure VI.10 are more likely to be the M_2 lunar semi-diurnal tide.

The composite lunar day analysis was also applied to the twenty-month data set available from the Rothera meteor radar. The results of this analysis for heights of 81-97 km are presented in Figure VI.13a, b. As with the Davis results, a clear wave-like feature is apparent with two cycles per lunar day, however, the amplitudes are

significantly larger, in this case reaching some 5 ms^{-1}. Comparisons of the amplitudes in the ~ 80 – 90 km height range where the two data sets overlap reveal amplitudes over Davis to be up to about 0.8 ms^{-1}, whereas the amplitudes over Rothera for the same height range reach up to about 3.0 ms^{-1}. Because the two data sets are of different lengths, an analysis of the Davis data was undertaken, using only the February 2005 to September 2006 interval when data was available from both radars simultaneously. This comparison (not shown) confirms that the amplitudes observed over Rothera are larger than those observed over Davis during this interval.

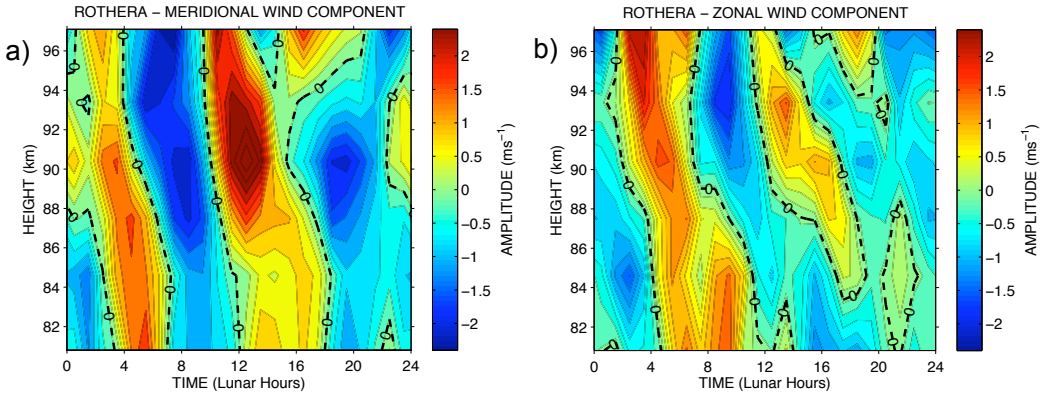


Figure VI.13a, b: Composite lunar day of the winds over Rothera for the interval 2005 to 2006, measured from the time of the first new Moon in this interval. a) zonal winds and b) meridional winds. The zero line is marked by the dashed line.

In order to determine the amplitude and phase of the lunar M_2 tide in each height gate, an analysis was again applied using the least-squares fitting method of Malin and Schlapp (1980). This method was used in the studies of Schlapp and Harris (1993); Stening *et al*. (1997b; 2003); Stening and Jacobi (2001); Niu *et al*. (2005); Sandford *et al*. (2006) and Sandford and Mitchell (2007a). The output of the analysis is thus monthly-mean estimates of the tidal phase and amplitude. When data from more than one year are considered, the monthly-mean values from each year are vector averaged to yield a mean. Note that in cases where there is significant inter-annual variability in the tidal phase this will lead to some degree of self cancellation and so reduce the derived tidal amplitudes (see Antarctic discussion Section VI.2.2).

To determine the errors in the amplitude and phase, the method of Malin and Schlapp (1980) is also followed. This method is used because when using the error on the fit to the combined tidal components, the resulting error will indicate the combined error to all of the fitted components. As the solar diurnal and semi-diurnal tides are often much larger, the error on the fit would therefore be biased toward those tides. A more appropriate error estimate has been shown by Malin and Schlapp (1980); in this method the hourly data points are randomly sorted into subsets. The solar diurnal S_1,

semi-diurnal S_2, ter-diurnal S_3, quarter-diurnal S_4 and the lunar semi-diurnal tide are then fitted to each subset using the method above to obtain the amplitude and phase of each oscillation fitted. The error on the amplitude and phase are then calculated from the standard deviation of these values.

VI.2.1 Antarctic Results

Amplitudes and phases for the lunar tide have been calculated on a monthly-mean basis throughout both the Rothera and Davis data sets. The monthly means of M_2 tidal amplitude and phase were then vector averaged to reveal the seasonal behaviour of the tide. Figure VI.14 presents the amplitude and phase of the tide over Rothera for both the zonal and meridional components, averaged over the height range 78 – 92 km. This height range was chosen because a comparative height range is available in both data sets and model results presented by Vial and Forbes (1994), therefore, allowing direct comparison. The model results are also shown in the figures.

The monthly-mean amplitudes over Rothera, averaged over heights of 78 – 92 km, are shown in Figure VI.14a. It can be seen that the amplitudes range from ~ 1 ms^{-1} to about 5 ms^{-1} and the zonal and meridional amplitudes are approximately similar. Further, there is a clear seasonal cycle with a maximum in summer (December to February), equinoctial minima and a second, smaller, maximum in winter (May to June). The model results are also indicated in the figure and, although they exhibit the observed seasonal cycle, the model amplitudes are significantly smaller than those observed around the summer and winter maxima.

The monthly-mean phases over Rothera, again averaged over heights of 78 – 92 km, are shown in Figure VI.14b. The observations reveal a simple seasonal cycle in which phases decrease in winter by about 6 – 8 lunar hours from their summer values. The zonal and meridional components are different by about 2.5 lunar hours with the zonal component leading. Since the zonal and meridional amplitudes are very close, this corresponds to an approximate circularly polarised rotating tidal wind vector that rotates with time in an anticlockwise sense when viewed from above. The model phases are also shown and their agreement with the observations is generally good. The observed seasonal cycle is well reproduced by the model. The mean variation between the observations and the model is ~ 1.2 lunar hours. The largest differences between the observations and model occur in late spring (September to November) where the model leads the observations by ~ 5 lunar hours.

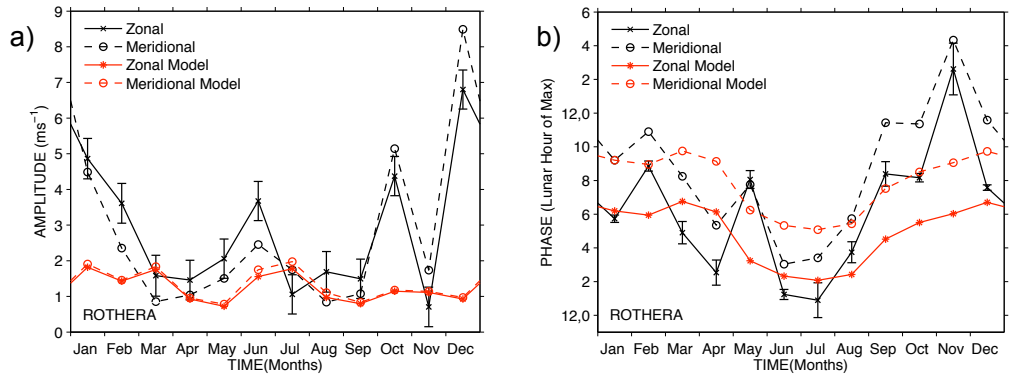


Figure VI.14a, b: Monthly-mean a) amplitudes and b) phases of the lunar M_2 tide over Rothera in the height range 78 – 92 km (black lines), solid lines indicate the zonal and dashed the meridional. Vector averaging is used. The time interval is February 2005 to September 2006. Error bars indicate the standard deviation as described above (similar for both components and therefore only marked on the zonal component for clarity). Monthly-mean results from model of Vial and Forbes (1994) are indicated by the red lines.

A similar analysis applied to the Davis data for the height range 77 – 91 km is presented in Figure VI.15. This does not use the full height range accessible to the radar, but is chosen so as to make the data comparable to that recorded over Rothera. Further, the analysis is restricted to only data recorded over the same interval of time as that from which data is available from both stations, i.e., February 2005 – September 2006. This is to allow a direct comparison with the data recorded over Rothera. The figure is also comparable with that of the UK (Figure V.11a, b), Esrange (Figure V.13a, b) and Ascension Island (Figure VI.2a, b).

Considering the amplitudes over Davis presented in Figure VI.15a, it can be seen that they are smaller than those observed over Rothera in most months. As an average over the year, the amplitudes over Davis are some 58% of those over Rothera. The very conspicuous summer-time maxima observed over Rothera in December – February is only weakly present over Davis, although there is a suggestion of a secondary maximum in winter as suggested by the model observed over Rothera. The phases over Davis are presented in Figure VI.15b. The seasonal variation in phase observed over Rothera is also apparent over Davis, with the phases moving to earlier values in winter. In comparison to the model, the phases are systematically later by a mean of about 4.5 lunar hours throughout the year.

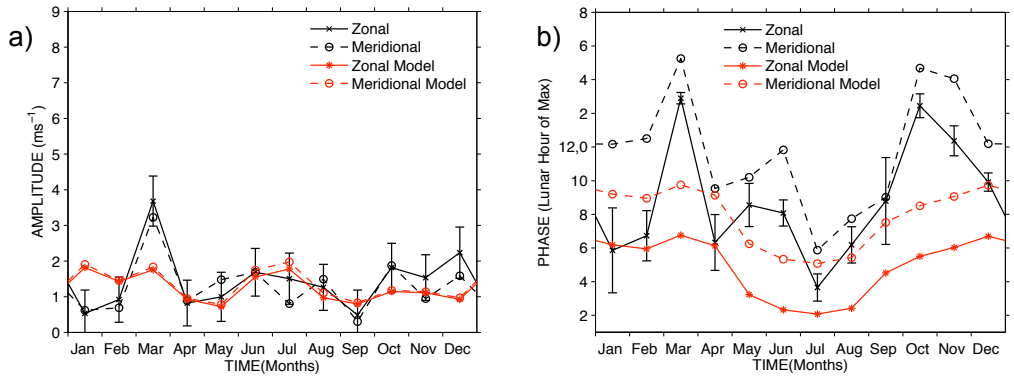


Figure VI.15a, b: Monthly-mean a) amplitudes and b) phases of the lunar M_2 tide over Davis in the height range 77 – 91 km (black lines), solid lines indicate the zonal and dashed the meridional. Vector averaging is again used. The time interval is February 2005 to September 2006. Error bars indicate the standard deviation as above (only marked for the zonal component for clarity). Monthly-mean results from model of Vial and Forbes (1994) are indicated by the red lines.

The full data set from Davis, however, spans some 13 years. An analysis identical to that above was applied to this entire dataset. The results are presented in Figure VI.16a, b. It is immediately obvious that the 13-year mean amplitudes are reduced over those observed in the ~ twenty-month dataset used for Figure VI.15a, b above. The amplitudes are significantly smaller, always being below 1.5 ms^{-1} and usually less than 1.0 ms^{-1}. The observed amplitudes have a seasonal cycle with a winter maximum of ~ 1.5 ms^{-1} and a summer minimum of ~ 0.5 ms^{-1}. In contrast, the model results suggest that there should be a semi-annual variation in amplitude, with maxima in winter (June - July) and late summer (January - March). Generally, the observed amplitudes are rather smaller than the model predictions and the main difference in the seasonal cycle is that the observations do not reveal the summer-time maximum of the model (January - March).

The significant differences between the amplitudes of the twenty-month analysis (Figure VI.16a) and the 13-year analysis (Figure VI.15a) appear to result from a high degree of inter-annual variability in the tidal phase. In the vector averaging used to construct the means this leads to a self-cancellation of the amplitudes. An analysis (not shown) using geometric averaging rather than vector averaging yields significantly higher mean amplitudes, generally of about 3 ms^{-1}.

The corresponding monthly-mean phases observed over Davis, calculated as vector means over the 13-year dataset, are shown in Figure VI.16b. As with Rothera, the zonal phases lead the meridional by ~ 3 lunar hours. The zonal and meridional amplitudes are again very similar, so in combination with the ~ 90º phase difference between the two components, this means that the tide is approximately circularly

polarised, consisting of a constant speed wind vector that rotates anti-clockwise with time. The phases display a clear seasonal cycle, very similar to that over Rothera, in which phases decrease in winter by about 6 lunar hours from their summer values. This seasonal cycle is also reproduced quite well by the model. The major difference between these observations and those made over Rothera is that here the observed phases in both components systematically lead the model phases by a mean of 5.8 lunar hours. These phases also lead those observed over Rothera by ~ 4.6 lunar hours (i.e., they are in approximate anti-phase).

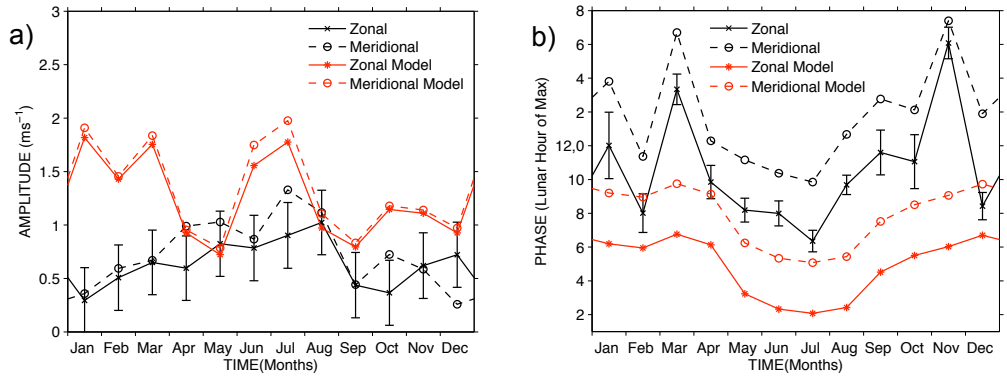


Figure VI.16a, b: Monthly-mean a) amplitudes and b) phases of the lunar M_2 tide over Davis in the height range 78 – 92 km (black lines), solid lines indicate the zonal and dashed the meridional. Vector averaging is used. The time interval is May 1994 to September 2006. Error bars marked on the zonal component. Monthly-mean results from model of Vial and Forbes (1994) are indicated by the red lines.

In summary, the lunar tide over Rothera appears to have a number of differences from that observed over Davis. In particular:

1. Amplitudes over Rothera are significantly larger (when considering vector means); the difference amounting to a factor of ~ 1.7 when averaged over the twenty-month interval of simultaneous observations.

2. The seasonal cycle of amplitudes over Rothera displays a clear semi-annual variation with a summer-time maximum and winter-time secondary maximum, whereas over Davis, during the same twenty-month interval, the seasonal cycle appears less well defined. In fact, in the 13-year mean analysis the largest amplitudes are observed over Davis in winter.

3. The phases over both sites vary seasonally in a manner that agrees with the model, i.e., the phases have their latest values in summer and decrease by about 6 lunar hours to reach the earliest values around midwinter. However, the phases over Davis

are systematically different by ~ 5 lunar hours from the model and those observed over Rothera.

The vertical structure of the M_2 lunar tide over the two sites was investigated by calculating the tidal amplitude as a function of height for each month. The data from a particular month in all the years available were then averaged to yield monthly means based on two years of data for Rothera and thirteen years of data for Davis. Figure VI.17a, b present contours of the amplitude of the zonal and meridional components of the lunar M_2 tide over Rothera and Figure VI.18a, b for Davis.

Both sets of contours of Figure VI.17 reveal a generally similar and well-defined seasonal behaviour in which the largest amplitudes of ~ 8 ms^{-1} are reached in winter (December - January) between 84 and 90 km, with the secondary maxima around June – July and a localised maximum in October maximising around 82 km. For the rest of the time, amplitudes are generally small (< 2 ms^{-1}). The contours of Figure VI.18 from Davis reveal more complicated and less well-defined behaviour; the amplitudes are much smaller, usually less than < 2 ms^{-1}, however, the height range of this radar is lower. Amplitudes tend to increase with height except for a maximum low down < 72 km in both components during April – May.

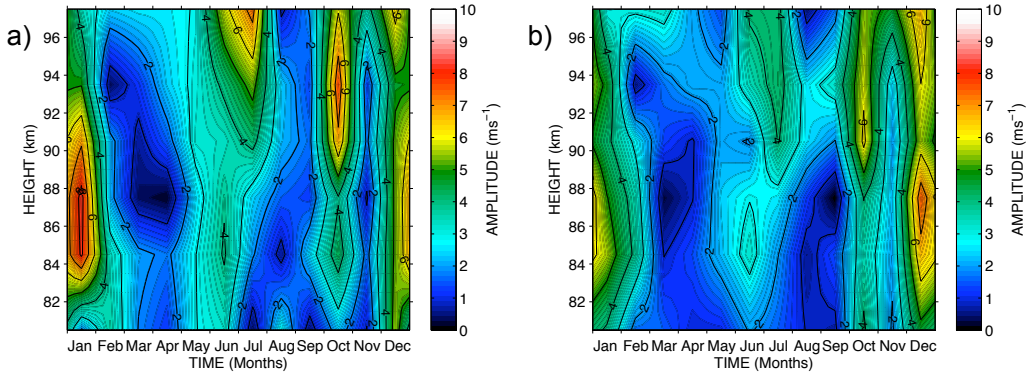


Figure VI.17a, b. Contour plot of the Lunar M_2 tidal amplitude calculated from monthly-mean amplitudes (vector averaged to give the seasonal variation) of a) zonal and b) meridional wind data over the period 2005 – 2007 from the Rothera meteor radar.

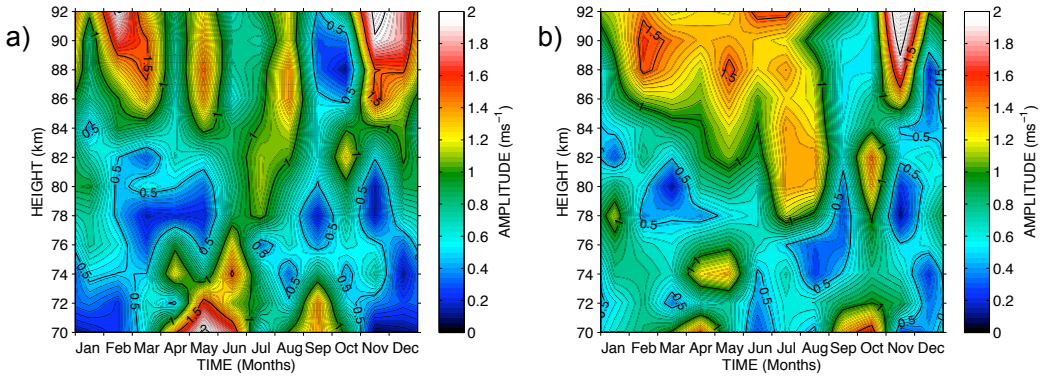


Figure VI.18a, b. Contour plot of the Lunar M_2 tidal amplitude calculated from monthly-mean amplitudes (vector averaged to give the seasonal variation) of a) zonal and b) meridional wind data over the period 1994 – 2006 from the Davis MF radar.

To compare this behaviour with that of the solar S_2 semi-diurnal tide, the same analysis was applied. Figures VI.19a, b and VI.20a, b present amplitude contours of the a) zonal and b) meridional components of the solar S_2 tide for Rothera and Davis, respectively. As before, the data have been vector averaged to produce a composite year.

Both sets of contours in Figures VI.19 and VI.20, reveal very similar behaviour to each other. Maxima are seen in both components at both stations during March to June and August to September. Over Rothera there is a distinct minimum in November, whereas over Davis at the same time, there is a maximum. Amplitudes tend to increase with height. Even though the height range is lower, the amplitudes over Davis are again lower than those over Rothera. This is possibly due to instrument bias. The amplitudes of these figures are also distinctly different from those of the lunar M_2 tide.

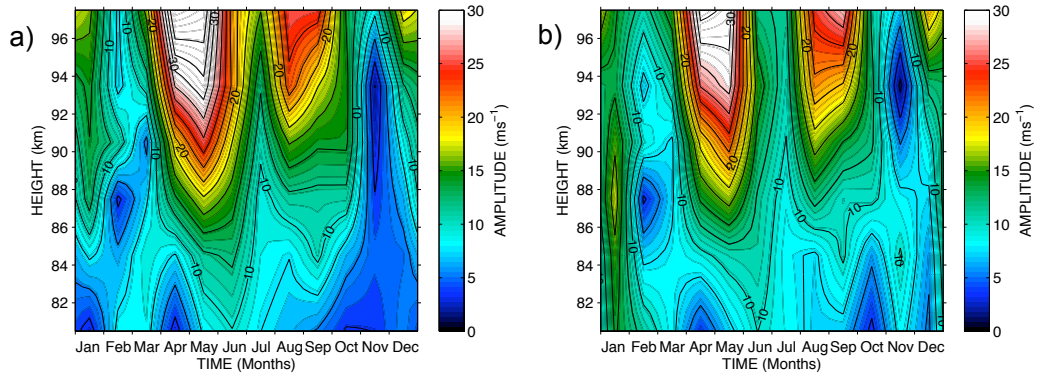


Figure VI.19a, b: Contour plot of the Solar S_2 tidal amplitude calculated from monthly-mean amplitudes (vector averaged to give the seasonal variation) of a) zonal and b) meridional wind data over the period 2005 – 2007 from the Rothera meteor radar.

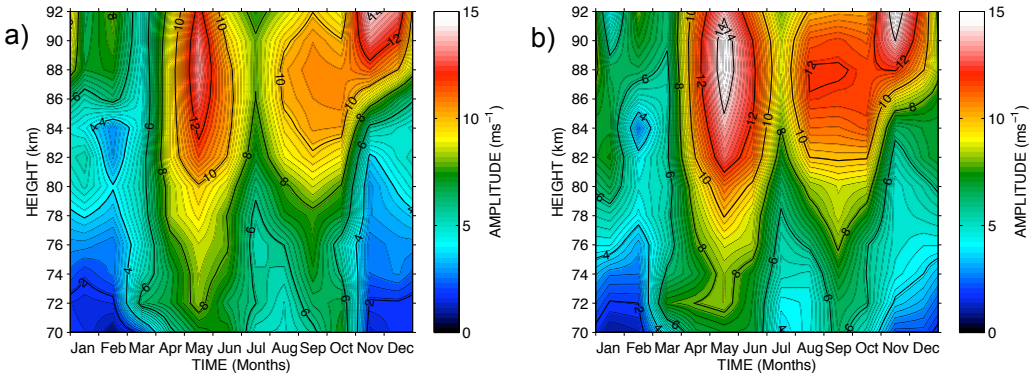


Figure VI.20a, b: Contour plot of the Solar S_2 tidal amplitude calculated from monthly-mean amplitudes (vector averaged to give the seasonal variation) of a) zonal and b) meridional wind data over the period 1994 – 2006 from the Davis MF radar.

The lunar M_2 tidal results for amplitude and phase from the Antarctic are now compared with the Vial and Forbes (1994) model. Figures VI.21 and VI.22 present the averaged monthly-mean amplitudes and phases as a function of height, respectively, measured over Rothera. The predictions of the Vial and Forbes model are also indicated for comparison. Considering the amplitudes, a noticeable feature of the model is that in many months the amplitude is relatively constant up to about 85 km, above which it increases rapidly (e.g., May – September and some other months). In contrast, the observations suggest that the amplitude often remains relatively constant across the ~ 80 – 97 km height range observed. This has the consequence that at heights below ~ 90 km, the observed amplitudes are usually either about equal to or larger than those of the model, whereas at heights above ~ 90 km, the model amplitudes exceed those observed (e.g., March – September).

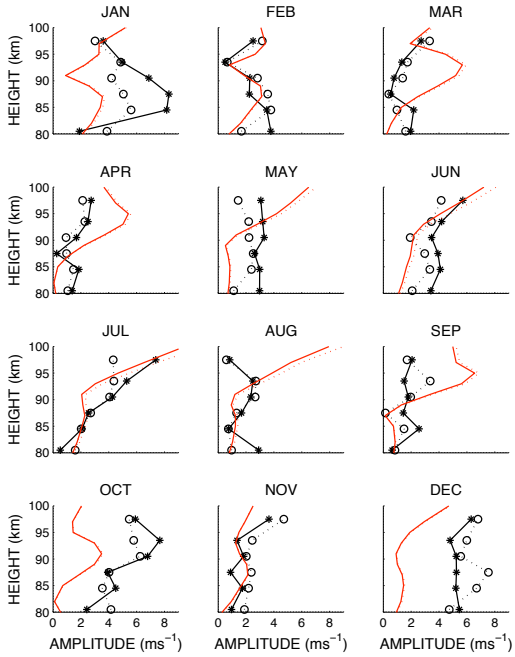


Figure VI.21: Monthly-mean amplitude vs. height of the lunar M_2 tide over Rothera, vector averaged from February 2005 to December 2006. The zonal components are solid lines marked by stars and the meridional, dashed lines marked by circles. Also shown are the results from the Vial and Forbes (1994) model, zonal results solid lines, meridional results dashed lines.

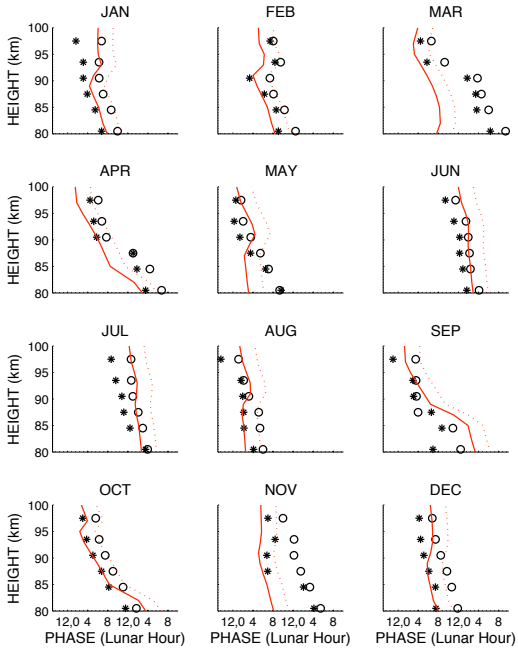


Figure VI.22: Monthly-mean phase vs. height of the lunar M_2 tide over Rothera, vector averaged from February 2005 to December 2006. The zonal components are marked by stars and the meridional by circles. Also shown are the results from the Vial and Forbes (1994) model, zonal results solid lines, meridional results dashed lines.

The phases of Figure VI.22 show that in nearly all months the observed phase over Rothera decreases with height and that the zonal component leads the meridional. The vertical phase gradient allows estimation of the vertical wavelength. If this is calculated using the entire 80 – 97 km height range then the results for the vertical wavelength are as indicated in Figure VI.23. It can be seen from the figure that there is a semi-annual cycle in the vertical wavelengths. They are longest in summer and winter (typically ~ 38 km) and shortest around the equinoxes (typically 15 km). In more detail, the observed vertical wavelengths vary from values as short as about 10 km in April to values as long as about 66 km in June. The vertical wavelengths calculated using the model data from the same height range are also indicated on the figure. In some months there is excellent agreement between the observations and model, particularly around the equinoxes. However, in other months the model wavelengths are significantly longer than those observed. The differences are particularly large around mid-winter (June to August) when the model suggests vertical wavelengths can be as long as 155 km, c.f. the observed wavelengths of ~ 25-50 km.

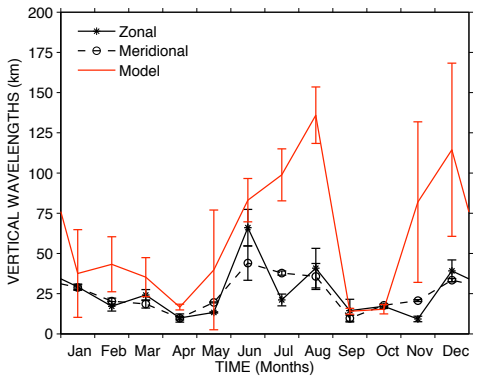


Figure VI.23: Vertical wavelength of the lunar M_2 tide at Rothera, calculated from vector averaged monthly means over the height range 78 – 100 km between February 2005 and December 2006. The zonal component is marked by the black solid line with stars, meridional by the black dashed line with circles and the average zonal and meridional results of the Vial and Forbes (1994) model by the red line.

A similar analysis to the above, applied to the Davis data is presented in Figures VI.24, VI.25 and VI.26. The vertical profiles of monthly-mean amplitudes are shown in Figure VI.24. The amplitudes are in general smaller than those observed over Rothera. As with the Rothera results, there are a number of months where the model suggests that the amplitudes should grow rapidly with height above about 85 km but where this is not observed to happen (e.g., January, February, March, April, September and October). As with the Rothera results, this means that the amplitudes agree reasonably well with the model up to about 85 km, but above that height the model amplitudes often exceed those observed.

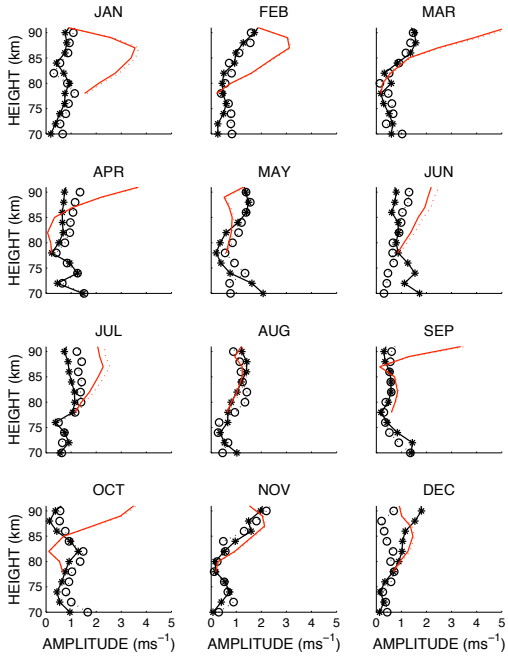


Figure VI.24: Monthly-mean amplitude vs. height of the lunar M_2 tide over Davis, vector averaged from May 1994 to September 2006. The zonal components are solid lines marked by stars and the meridional, dashed lines marked by circles. Also shown are the results from the Vial and Forbes (1994) model, zonal results solid lines, meridional results dashed lines.

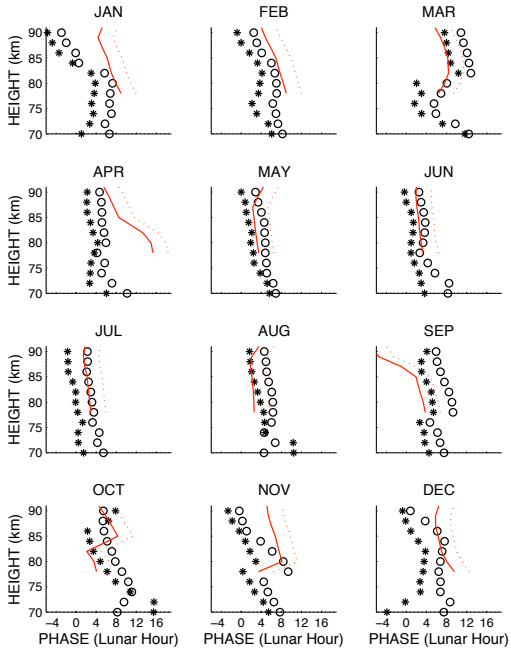


Figure VI.25: Monthly-mean phase vs. height of the lunar M_2 tide over Davis, vector averaged from May 1994 to September 2006. The zonal components are marked by stars and the meridional by circles. Also shown are the results from the Vial and Forbes (1994) model, zonal results solid lines, meridional results dashed lines.

The phases of Figure VI.25 decrease with height in nearly all months and the zonal component leads the meridional by approximately 3 lunar hours. The monthly-mean vertical wavelengths calculated using data from the 77 – 91 km height range are shown in Figure VI.26. This height range is selected so that the results are directly comparable with the available model data. From the figure, it can be seen that the observed vertical wavelengths are usually shortest in summer, with values of ~ 20 km, somewhat longer around the equinoxes, where wavelengths are ~ 58 km, and longest in winter, where they reach values as long as 65 km. A similar seasonal cycle is apparent in the model data, but the predicted winter-time vertical wavelengths are somewhat longer than those observed.

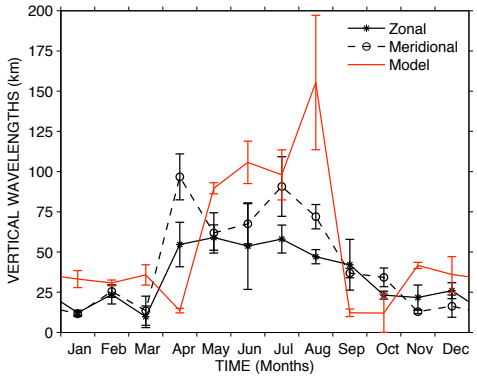


Figure VI.26: Vertical wavelength of the lunar M_2 tide at Davis, calculated from vector averaged monthly means over the height range 77 – 91 km between May 1994 and September 2006. The zonal component is marked by the black solid line with stars, meridional by the black dashed line with circles and the average zonal and meridional results of the Vial and Forbes (1994) model by the red line.

The Rothera and Davis stations are separated by only 1.1° in latitude, but by some 146° in longitude. This longitudinal spacing can be used to investigate the longitudinal structure of the lunar tide, similar to the analysis of Baumgaertner *et al.* (2006). For instance, if the lunar tide were a pure migrating M_2 mode, i.e., a westward, "Moon-following" mode of zonal wavenumber 2 observed by two stations spaced in longitude by 180°, then each station would see the tidal variation to be in phase (i.e., the lunar hour of maximum would be the same and so the phase difference would be 12 or equivalently 0 lunar hours). Departures from this would indicate the presence of non-migrating lunar tidal modes. For example, a hypothetical wavenumber 1 lunar semi-diurnal tide would produce a lunar semi-diurnal oscillation over each site, but in this case they would be in anti-phase (i.e., 6 lunar hour difference in phase for a semi-diurnal lunar tide). An example of the wavenumber 1 and 2 is shown in Figure VI.27. The argument can be extended to show that all even zonal wavenumbers (2, 4, 6…) would yield in-phase oscillations over the two sites and all odd zonal wavenumbers (1, 3, 5…) would yield anti-phase oscillations.

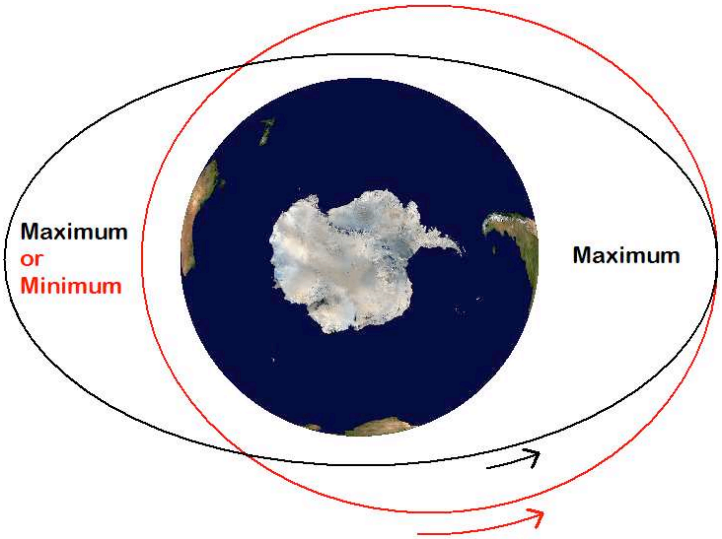


Figure VI.27: Wave diagram around Antarctica, the black ellipse indicates a wavenumber 2 wave with maxima occurring simultaneously over opposite sides of the pole and red circle indicates a wavenumber 1 wave with maxima occurring on one side of the pole as a minimum occurs on the other. For the lunar M_2 tide a westward propagating wavenumber 2 would be the migrating mode and to receive the same signature over a single station, the wavenumber 1 would need a phase speed which is twice as fast.

However, the stations are actually 146º apart in longitude, not 180º. This results in a difference in phase of 9.733 lunar hours, not 12 lunar hours between the sites, in the case of a migrating zonal wavenumber 2 lunar M_2 semi-diurnal tide. In this case the phase at Davis will lead that at Rothera. A hypothetical wavenumber 1, non-migrating, westward-travelling, lunar semi-diurnal tide in this circumstance would yield a difference in phase of 4.867 lunar hours, not 6 lunar hours, with Davis again leading Rothera. Similarly, a hypothetical wavenumber 1, non-migrating, eastward-travelling, lunar semi-diurnal tide would yield a difference in phase of 7.133 lunar hours, with Davis leading Rothera, travelling in the opposite direction. The relative phase between Rothera and Davis thus provides an insight into the zonal structure of the tide.

Figure VI.28a, b displays the difference in monthly-mean phase, calculated as Rothera minus Davis, over the height range 87 – 91 km for the zonal and meridional components respectively. The horizontal lines on each figure indicate the 9.733, 4.867 and 7.133 lunar hour difference in phase between Rothera and Davis that correspond to a westward wavenumber 2 (migrating), westward wavenumber 1 (non-migrating) and eastward wavenumber 1 (non-migrating) lunar semi-diurnal tide, respectively.

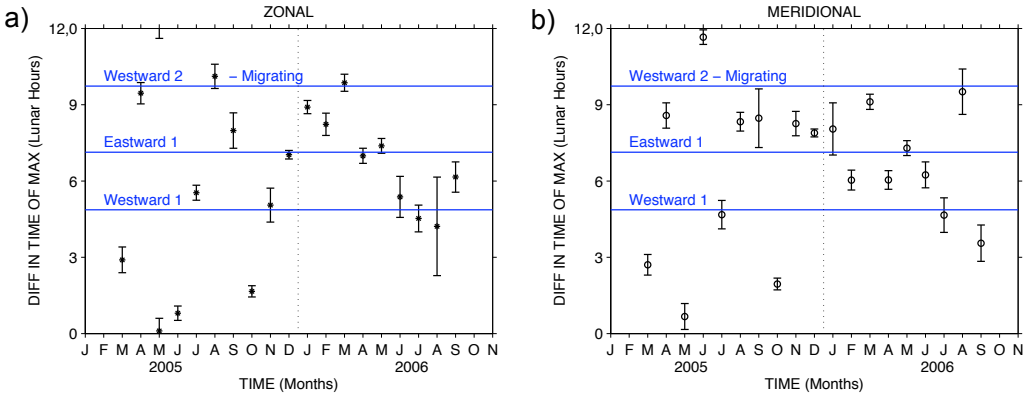


Figure VI.28a, b: Differences in the time of maximum of the lunar M_2 tide between Rothera and Davis (Rothera minus Davis) at 87-91 km for the a) zonal and b) meridional components calculated as monthly means between February 2005 and September 2006. Horizontal blue lines indicate the migrating, westward travelling, wavenumber 2 mode, non-migrating, westward 1 mode and non-migrating, eastward 1 mode. Error bars indicate the standard deviation.

As can be seen from the figure, only in a few months are the phase differences close to the 9.733 lunar hours that would be expected for a purely migrating lunar semi-diurnal tide. Most of the points in each figure fall within the range of about 4- to 10 lunar hours. There is no obvious tendency for the points to cluster near the lines corresponding to the eastward 1 or westward 1 non-migrating modes. To examine the difference in phase between Rothera and Davis in more detail, Figure VI.29 presents a histogram showing the distribution of the differences. It can be seen that most of the differences lie between 4 and 10 lunar hours. Smaller or larger differences are less common.

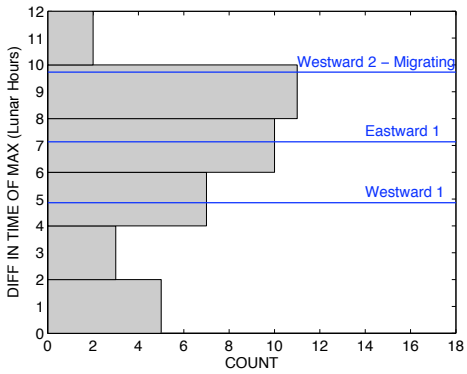


Figure VI.29: Histogram of the difference in time of maximum between Rothera and Davis (Rothera minus Davis) at 87 – 91 km, both the zonal and meridional components are separated into 2 hourly bins. Horizontal blue lines indicate the migrating, westward travelling, wavenumber 2 mode, non-migrating, eastward 1 mode and non-migrating, westward 1 mode.

This behaviour implies that the tide is rarely composed of a purely migrating wavenumber 2 mode and that the behaviour is more usually like that of a mixture of

migrating and non-migrating modes. Of course, it is possible that other, higher wavenumber, non-migrating modes are also present.

Figure VI.30 displays the same analysis applied to the solar S_2 tide. The figure shows the difference in monthly-mean phase, calculated as Rothera minus Davis, over the height range 87 – 91 km for the zonal and meridional components respectively. The horizontal lines on each figure again indicate the 9.733, 4.867 and 7.133 solar hour difference in phase between Rothera and Davis. These correspond to the phase differences expected from a westward wavenumber 2 (migrating), westward wavenumber 1 (non-migrating) and eastward wavenumber 1 (non-migrating) solar semi-diurnal tide, respectively.

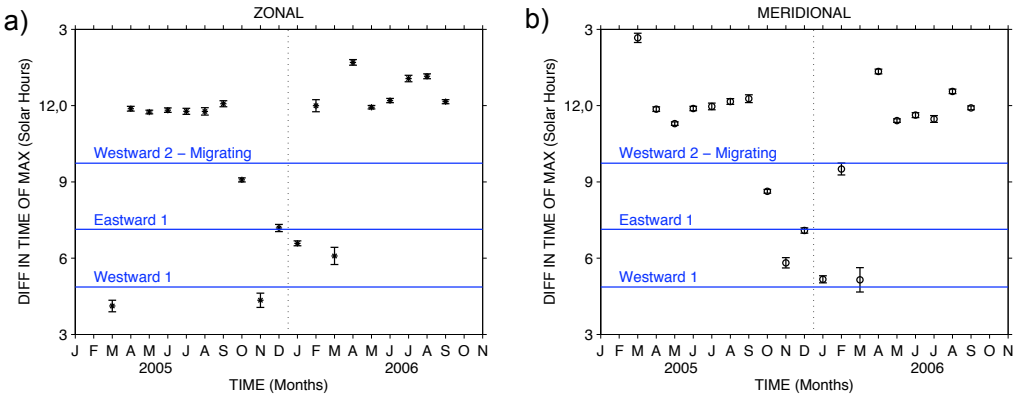


Figure VI.30a, b: Differences in the time of maximum of the solar S_2 tide between Rothera and Davis (Rothera minus Davis) at 87-91 km for the a) zonal and b) meridional components calculated as monthly means between February 2005 and September 2006. Horizontal blue lines indicate the migrating, westward travelling, wavenumber 2 mode, non-migrating, westward 1 mode and non-migrating, eastward 1 mode. Error bars indicate the standard deviation.

From the figure, it can be seen that a large number of the points lie on a line which is approximately 12 hours / 0 hours phase difference, especially during the extended winter seasons (April – September). This indicates that the solar tide is maximising at approximately the same time over both sites during these periods.

VI.2.2 Discussion of Antarctic Results

A small number of studies have examined the lunar M_2 tide in the middle- and high-latitude southern hemisphere MLT region. These studies include observations made over Christchurch (44°S, 173°E) and Scott Base (78°S, 166°E) by Stening *et al*. (1995). The case presented here thus addresses a latitude in between these two sites.

Stening *et al*. (1995) reported MF-radar observations made over Christchurch (44°S, 173°E) and Scott Base (78S, 167°E) during the interval 1983 – 1984 at heights of 82 – 100 km. They measured both the solar and lunar semi-diurnal tides, but noted some difficulties in isolating the lunar semi-diurnal tide. Over Christchurch, they found an annual seasonal cycle with maximum amplitudes occurring in winter, in contrast to the semi-annual seasonal cycle observed over Rothera but rather like the annual cycle observed over Davis. The monthly-mean tidal amplitudes over Christchurch reached values as large as 8 ms^{-1}, roughly comparable to those over Rothera (note that Stening *et al*. (1995) determined that lunar tidal amplitudes vary with local time. Their results presented the largest amplitudes and so may be overestimates compared to the analysis used here). Over Scott Base the observed seasonal behaviour had an annual cycle with largest monthly-mean amplitudes in winter.

Measurements of the lunar M_2 tide made at the conjugate latitude in the northern hemisphere (Esrange, 68°N, 21°E) over the interval 1999 – 2007 are presented in Chapter V. The meteor radars at Rothera and Esrange are identical instruments, so comparisons can be made with a minimum of instrument bias. Further, the data analysis is also identical. The seasonal cycle of amplitudes appears rather different over the two sites. Over Rothera the amplitudes are largest in summer, but with a secondary maximum in winter. In contrast, over Esrange the largest amplitudes are observed in winter with a secondary maximum in autumn. However, we should note that the level of inter-annual variability is high and so the differences may not be significant. The tidal phases and vertical wavelengths observed over Rothera and Esrange are generally in good agreement (when the phases over one site are shifted by six months so that similar seasons are compared).

Comparisons of the solar semi-diurnal tide indicate that the amplitudes over Davis are smaller than those seen over Rothera; this is likely to be due to instrument biases of the different radar techniques. Also the solar tide over Esrange had the same behaviour as that of the lunar tide, however, over Rothera the behaviour of the solar and lunar tides are significantly different. A fundamental difference between these tides is that the excitation of the solar tides is highly variable and time dependant on varying atmospheric concentrations, as described in Chapter V. The forcing by the Moon's gravitational field occurs predominantly at the lower boundary and in the denser, lower atmosphere. This forcing is well specified and virtually constant, therefore any variability observed in the MLT-region lunar tide will be primarily due to changes in the propagation characteristics of the tide as it ascends through the atmosphere from its source region. This indicates that in the northern hemisphere the tides are equally affected as they propagate, and the solar tides are likely to be less

affected by their varying source regions. However, in the southern hemisphere, the source variability is likely to be a major factor in the solar tidal variability seen in the MLT region.

We now consider the differences in phase between Rothera and Davis. The fact that the measured phase differences are variable and usually only a few lunar hours different from those that would be expected for a purely migrating tide suggests that the observed tide includes significant contributions from non-migrating modes. The analysis procedure used here is very sensitive in terms of frequency and selects only frequencies very close to that of the lunar semi-diurnal tide (as evidenced by the fact that it can discriminate between the lunar and solar semi-diurnal tides). This means that any non-migrating modes detected must also be of or very close to the lunar semi-diurnal frequency. Such modes could arise from a non-linear interaction between the migrating lunar tide and a stationary planetary wave (i.e., one with a frequency of zero). The "sum" and "difference" secondary wave produced in such an interaction would thus be of lunar semi-diurnal frequency. Given that stationary planetary waves usually have zonal wavenumber 1 or 2, this would produce non-migrating semi-diurnal lunar tides with wavenumbers of 1 (eastward), 1 (westward) or 3 (westward) in the case of a zonal wavenumber 1 planetary wave or 4 (westward) or zero in the case of a zonal wavenumber 2 planetary wave.

Finally, we should note that any "difference" waves produced by a non-linear interaction between the 12-hour solar tide and the 16-day planetary wave will be largely removed from the multiple-year superposed epoch analysis of The Antarctic case section. This is because they will be randomly phased with respect to the New Moon position in successive years and will thus self cancel. However, in consideration of individual monthly means, such waves may be present and introduce some variability.

VI.2.3 Conclusions of the Antarctic Studies

The lunar M_2 tide at Antarctic latitudes has been measured by meteor and MF radar over Rothera (68°S, 68°W) during 2005-2006 and Davis (68°S, 78°E) during 1994-2006. A clear signal corresponding to the M_2 lunar semi-diurnal tide is evident in the data from both locations and an analysis is applied that yields monthly-mean values of tidal amplitude and phase. The tide is found to have similar zonal and meridional amplitudes with the zonal component leading in phase by about 3 lunar hours – corresponding to an approximate circular polarisation, with an anticlockwise rotating tidal wind vector and an upward propagation of the tide.

Individual monthly-mean amplitudes are observed to reach values as large as 8 ms^{-1} and vary seasonally. Averaged over the above intervals of observation, over Rothera a semi-annual cycle in the amplitude is apparent with maxima in winter and summer, whereas over Davis the amplitudes are largest in winter. Amplitudes at the same heights are generally larger over Rothera than over Davis. However, the amplitudes and phases show a high degree of inter-annual variability, so it is not clear if these differences are significant. The situation is further complicated by possible measurement biases between the two techniques, as this behaviour is also seen in the solar S_2 tide over the two sites. In general, these biases may lead to amplitudes determined by meteor radar being larger than those determined by MF radar which was also observed by Manson *et al*. (2004). Vertical wavelengths are generally quite short, typically 15 km around the equinoxes, increasing to about 70 km in winter.

Zonally-averaged results from the Vial and Forbes (1994) model reproduce the observed seasonal behaviour of the amplitudes quite well over both sites. However, there are systematic and large differences in phase between the model and the observations over Davis.

Comparisons of the relative phase of the tide between Rothera and Davis suggest that it cannot consist of a purely migrating zonal wavenumber 2 lunar tide and may well contain contributions from non-migrating lunar tidal modes.

Note that the work presented in this Chapter was published as: Sandford and Mitchell, 2007a and Sandford *et al.*, 2007b.

Chapter VII

VII. Conclusions

VII.1 Arctic Planetary Waves

The work presented in this thesis aims to address a puzzle found in recent observations of the polar mesosphere. Specifically, it has been discovered that waves with periods near two days reach significant amplitudes in both summer and winter. This is in striking contrast to mid-latitude observations where two-day waves maximise in summer only. Data has been used from a meteor radar at Esrange (68ºN, 21ºE) in the Arctic and data from the MLS instrument aboard the EOS Aura satellite to investigate the wintertime polar two-day wave in the stratosphere, mesosphere and lower thermosphere. The radar data reveal that mesospheric two-day wave activity measured by horizontal-wind variance has a semi-annual cycle with maxima in winter and summer and equinoctial minima. The MLS data reveal that the summertime wave in the mesosphere is dominated by a westward-travelling zonal wavenumber 3 wave with significant westward wavenumber 4 present. It reaches largest amplitudes at mid-latitudes in the southern hemisphere. However, in the winter polar mesosphere, the wave appears to be an eastward-travelling zonal wavenumber 1 and 2, which is not seen during the summer. At the latitude of Esrange, the eastward 2 wave reaches maximum amplitudes near the stratopause and appears related to similar waves previously observed in the polar stratosphere. After careful consideration of possible excitation sources, it is concluded that the wintertime polar two-day wave is most likely the mesospheric manifestation of an eastward-propagating wave originating in the stratosphere, maximising at the stratopause and likely to be generated by instabilities in the polar night jet.

VII.2 The Lunar Semi-Diurnal Tide

The Lunar M_2 tide has been investigated in the region of the mesosphere and lower thermosphere (~ 80 – 100 km) over a number of different sites using meteor wind radar and medium frequency radar. An analysis technique has been developed to extract the amplitude and phase of the lunar M_2 tide from the spectrum of winds over each location.

These studies have identified features of the wind field which have previously been unobserved, including: observations of non-migrating lunar tides and a beating between the solar S_2 and the lunar M_2 tides.

A beating between the solar S_2 and the lunar M_2 tides could produce an apparent modulation of the S_2 tide at periods near 16 days. Close inspection is therefore required to avoid misinterpreting a lunar tidal beating as a non-linear interaction between the 16-day planetary wave and the solar S_2 tide.

The lunar M_2 tide has been observed in both hemispheres, at almost identical latitudes and using almost identical instruments. The tidal amplitudes were found to be very similar in the two hemispheres. However, there are striking differences found in similar observations of the solar S_2 tide. As both tides are at very similar frequencies and would have to propagate through the same atmosphere to reach the MLT region, this suggests that the way the tides propagate through the atmosphere between the two hemispheres is almost identical and that the differences observed are likely to come from the variable sources of the solar S_2 tide i.e., differing concentrations of ozone and oxygen between the two hemispheres.

Evidence has also been presented which indicates that there are likely tides which have the same period as the lunar M_2 tide, but are not Moon-following, i.e., non-migrating lunar tides, and these reach significant amplitudes in the MLT region.

VII.2.1 The Lunar M_2 Tide at Middle and High Arctic Latitudes

Meteor radar has been used to measure the horizontal winds in the MLT region over Castle Eaton (52°N) in the UK. These observations comprise the longest study of the lunar tide in the MLT region to date, covering an interval of 16 years.

The amplitudes of the M_2 tide were found to be largest in winter (up to ~ 11 ms^{-1}), with a secondary maximum in autumn. However, annual-mean amplitudes show that the tide exhibits significant year-to-year variability. The zonal and meridional phases are in quadrature, with the meridional component leading the zonal component and as both components have similar amplitudes this indicates that the tide is circularly polarised and rotating clockwise viewed from above. Comparisons with models and other observations near this latitude are in generally good agreement.

A second meteor radar at Esrange (68°N) in Arctic Sweden, with a 6-year data set has also been used to reveal the signature of the 12.420-hour lunar M_2 tide. These observations of the tide at Esrange make up the longest continuous data set at Arctic latitudes and are currently the most northerly study of the lunar tide.

Over Esrange, the tide is observed to reach very significant amplitudes as large as ~ 11 ms^{-1}. This radar has height finding capability and so allows the investigation of the vertical structure of the lunar tide, indicating that the sources of the M_2 tide are below this height region. Very similar to the UK, the tide has maximum amplitudes in winter with a second autumnal maximum, and the phases are found to be in quadrature and circularly polarised, rotating the same way as found over the UK. The amplitude is found to increase with height over the 80 - 100 km height range observed. Vertical wavelengths are very variable, ranging from about 16 km in summer to about 60 km in winter. Comparisons with the Vial and Forbes (1994) model and the Poker Flat (65°N) MST radar reveal generally good agreement. The only exception is in the case of summertime vertical wavelengths, which are observed to be significantly shorter than predicted by the model.

VII.2.2 The Lunar M_2 Tide at Equatorial and High Antarctic Latitudes

A meteor radar has been used to measure the horizontal winds in the equatorial mesosphere and lower thermosphere over Ascension Island (8.0°S, 14.4°W). A 5-year data set covering the interval 2001 – 2005 over the height range 78 – 100 km has been considered. The lunar M_2 tide is clearly evident in the data and reaches amplitudes as large as 11ms^{-1} in the meridional component and 6 ms^{-1} in the zonal component. These are the first observations of the lunar tide made over the equatorial Atlantic sector. Comparisons of the observed seasonal behaviour with the model of Vial and Forbes (1994) reveals good agreement, but the observed amplitudes are generally larger and there is a systematic phase difference of ~ 2 lunar hours with the observed phases lagging the model. Comparisons with observations made at other equatorial

sites suggest the presence of non-migrating lunar M_2 tides and/or significant inter-annual variability.

Measurements over Antarctica have also been used to observe the lunar M_2 tide. These include, a meteor radar at Rothera (68°S, 68°W) and MF radar at Davis (68°S, 78°E). Data from Rothera recorded over a 2 year interval in 2005 – 2006 and data from Davis recorded over the 13-year interval 1994 - 2006 were examined to investigate the monthly-mean behaviour of the tide. The amplitudes reach values as large as 8 ms^{-1}. The vertical wavelengths of the tide vary seasonally from ~ 10 – 65 km. Comparisons of the phase of the tide measured over the two sites reveal that it does not purely consist of a migrating wavenumber 2 mode. This suggests that other, non-migrating, modes are likely to be present.

VII.3 Suggestions for Future Work

The following suggestions are made as the logical continuation of the studies presented in this thesis:

1) The Source of the E1 planetary wave

Signatures of an eastward travelling, zonal wavenumber 1, planetary wave were found in the two-day wave studies. A similar E1 wave has been observed in the wintertime high-latitude stratosphere. This wave has been reported to have a four-day period and have the same source as the E2 planetary wave studied here. (i.e., instabilities in the four-day rotation of the polar night vortex). Using a similar satellite study as in the E2 wintertime two-day wave, this hypothesis could be confirmed or falsified.

2) Kelvin waves using MLS and Ascension Island meteor radar measurements

Also seen in the E1 planetary wave results is a large amount of activity over the equatorial region. This activity is likely to be ultra-fast Kelvin waves as these are eastward travelling equatorial waves. However, Rossby-gravity waves are also seen at similar periods. Using the results from the Ascension Island meteor radar, Kelvin waves can be distinguished from Rossby-gravity waves by looking for times when there are large zonal amplitude and no meridional amplitudes (since the Kelvin waves are zonally polarised). The geopotential height or temperature measurements from the Aura MLS satellite can then also be used to produce the first satellite climatology (i.e., direction and wavenumber analysis) of ultra-fast Kelvin waves and confirm that what is seen is actually eastward travelling to complement the radar results.

3) Global Quasi-2-Day Wave

A climatology of the quasi-two-day wave can be constructed using satellite data. In the literature there are a large number of studies of the quasi-two-day wave using single point measurements (e.g., radar), there are also a few collaborative campaign studies of the quasi-two-day wave using multiple stations. For example, the PSMOS campaign (Pancheva *et al*., 2004) which studied the quasi-two-day wave using 15 radars in the northern hemisphere to investigate its global structure. This analysis could now be complemented using satellite instrumentation. Using satellite data would avoid instrument bias. It would also give measurements on a synoptic scale which is not possible with individual stations.

4) Coordinated, Multi-Satellite Studies

Due to the orbit of Sun-synchronous satellites periods relating to the solar day cannot be measured. In the case of precessing satellites they require months of integration

time, which is only able to make long period seasonal and interannual climatologies. However, shorter-period variations, which occur on timescales of days, are difficult to measure globally. Observations from multiple satellite measurements, e.g., SABER onboard TIMED and MLS onboard Aura, which both measure temperatures, could be combined to create daily estimates of tidal properties.

5) Modelling study of the lunar M_2 tide

The next step in this analysis of the lunar tide would ideally involve modelling work to determine the causes of the structure seen over the locations presented. The lunar tide has previously been modelled by Vial and Forbes (1994). This is a zonally symmetric model of the M_2 tide, which predicts the features and general structure of the tide very well. However, it does not include any longitudinal information. Therefore, with more modern results the model could be updated and even be included in the Global Scale Wave Model (Hagan *et al.*, 1995), CMAT (Harris *et al.*, 2002) and other similar global tidal models.

6) Stratospheric Warmings

This would be a study of how the MLT region interacts with the underlying stratosphere. It has already been seen (Chapter IV) that stratospheric warmings have an effect on the dynamics of the MLT region through planetary-wave activity. However, the extent of these effects is still uncertain. Using the data from the UK Met Office, Stratospheric Assimilated Analysis, along with meteor wind radar measurements from both the Arctic and Antarctic regions, the effects can be studied. The Arctic has a weaker polar vortex and numerous major stratospheric warmings and the Antarctic has a strong polar vortex and only one observed major stratospheric warming in 2002.

7) Riogrande Radar Installation

The Riogrande radar is a new installation and an advanced SKYiMET meteor radar. It is specifically designed to look for gravity waves and therefore is ideal for the study of gravity wave/tidal interactions. It uses a much higher transmission power than the standard SKYiMET systems and is able to detect many more meteors. The radar is located on the southern tip of Argentina and the distribution of meteors collected is shown in Figure VII.1. The distribution of meteors from the Rothera meteor radar is also shown in the figure. These two stations provide a unique view of the polar vortex as, on average during winter, Rothera will be inside and Riogrande outside, with only a short distance between the two stations, as shown in Figure VII.2.

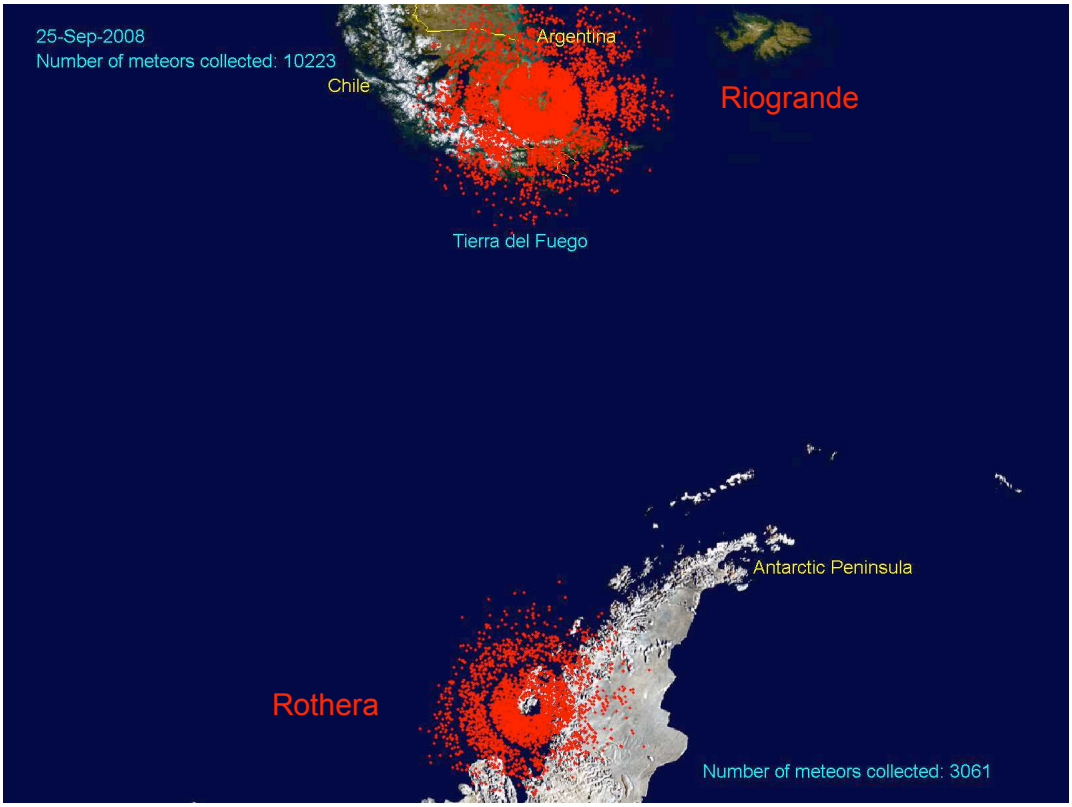


Figure VII.1: The meteor distributions of the Riogrande (Top) and Rothera (Bottom) Meteor Radars on the 25th September 2008. The Figure gives an idea of the distance between the two stations. For the majority of the time, the edge of the polar vortex will pass in between these two stations.

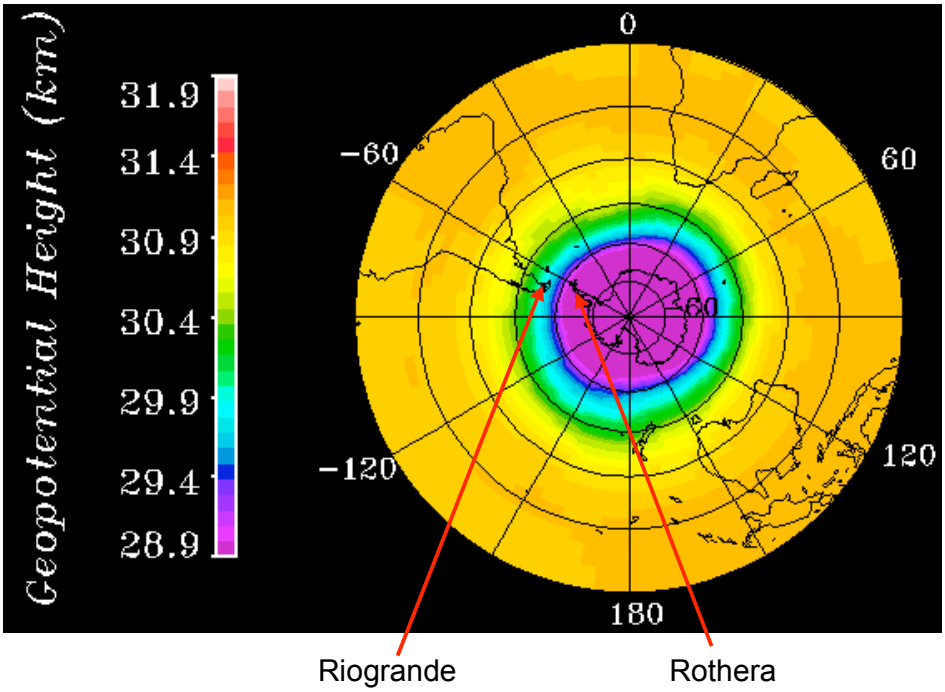


Figure VII.2: U.K.M.O. geopotential height measurements at the 10 hPa level on 15th July 1995. The Polar Vortex is indicated by low values surrounding the pole.

8) Bear Lake Installation

A meteor wind radar is to be installed at the Bear Lake Observatory (42°N), mid-latitude in the northern hemisphere. With the decommissioning of the UK meteor radar (55°N) and the planned collaboration to install a system in southern Argentina at Riogrande (53°S), this location would complete a latitudinal chain of radars and allow for a global study of MLT dynamics. The Bear Lake Observatory (42°N) houses a variety of complementary optical instruments managed by Utah State University. This will allow for collaborative imager-radar studies.

VII.4 Bibliography

Akmaev, R. A., Fomichev, V. I. and Zhu, X.: Impact of middle-atmospheric composition changes on greenhouse cooling in the upper atmosphere, *J. Atmos. Sol. Ter. Phys.,* vol. 68 pp. 1879–1889, 2006.

Andrews, D. G., Holton, J. R. and Leovy, C. B.: Middle Atmosphere Dynamics, Academic Press, 1987.

Baumgaertner, A. J. G., Jarvis, M. J., McDonald, A. J. and Fraser, G. J.: Observations of the wavenumber 1 and 2 components of the semi-diurnal tide over Antarctica*, J. Atmos. Terr. Phys.*, 68 (11), 1195-1214, 2006.

Beer, T.: Atmospheric Waves, Adam Hilger Ltd, London, 1974.

Beldon, C. L., Muller, H. G. and Mitchell, N. J.: The 8-hour tide in the mesosphere and lower thermosphere over the UK, 1988-2004, *J. Atmos. Sol.-Terr. Phys.*, 68, 655-668, 2006.

Beig, G., Keckhut, P., Lowe, R. P., Roble, R. G., Mlynczak, M. G., Scheer, J., Fomichev, V. I., Offermann, D., French, W. J. R., Shepherd, M. G., Semenov, A. I., Remsberg, E. E., She, C. Y., Lubken, F. J., Bremer, J., Clemesha, B.R., Stegman, J., Sigernes, F. and Fadnavis, S.: Review of mesospheric temperature trends, Rev. Geophys., 41, 4, 1015, 2003.

Benedick, R.E.: Greenhouse warming: Negotiating a global regime, World Resources Institute, 1991.

Brasseur, G. and Solomon, S.: Aeronomy of the Middle Atmosphere: Chemistry and Physics of the Stratosphere and Mesosphere, Third revised and Enlarged edition, Dordrecht – Springer, 2005.

Brekke, A.: Physics of the Upper polar atmosphere, John Wiley & Sons, 1997.

Bromberg, I.: Short-Term Variations of the Lunar Cycle, http://www.sym454.org/lunar/, University of Toronto, Canada, 2007.

Ceplecha, Z., Borovicka, J., Elford, W. G., Revelle, D. O., Hawkes, R. L., Porubcan, V. and Šimek, M.: Meteor phenomena and bodies, *Space Science Reviews* 84, 327-471, 1998.

Chapman, S. and Lindzen R. S.: Atmospheric Tides: Thermal and Gravitational, 201 pp., D. Reidel Publishing Company / Dorderecht – Holland, 1970.

Charney, J. G. and Drazin, P. G.: Propagation of Planetary-Scale Disturbances from the Lower into the Upper Atmosphere, *J. Geophys. Res.*, 66, 83-109, 1961.

Curtius, J., Weigel, R., Vossing, H. J., Wernli, H., Werner, A., Volk, C. M., Konopka, P., Krebsbach, M., Schiller, C., Roiger, A., Schlager, H., Dreiling, V., and Borrmann, S.,: Observations of meteoric material and implications for aerosol nucleation in the winter Arctic lower stratosphere derived from in situ particle measurements, *Atmos. Chem. and Phys.*, 5, 3053-3069, 2005.

Doodson, A. T.: The harmonic development of the tide-generating potential, *Proc. Roy. Soc. London*, A100, 305-329.

Dzubenko, N. I., Ivchenko, V. N. and Kozak, L. V.: Temperature variations in the thermosphere over the earthquake focuses as inferred from satellite data, *Geomag. And Aeron.*, 43, 118-123, 2003.

England, S. L., Dobbin, A., Harris, M. J., Arnold, N. F. and Aylward, A. D.: A Study into the effects of gravity wave activity on the diurnal tide and airglow emissions in the equatorial mesosphere and lower thermosphere using the Coupled Middle Atmsophere and Thermosphere (CMAT) general circulation model, *J. Atmos. Solar-Terr. Phys.*, 68, 293-308, 2006.

Forbes, J. M.: Tidal and Planetary Waves, in The Upper Mesosphere and Lower Thermosphere: A Review of Experiment and Theory*, Geophys. Monogr.*, vol. 87, edited by Johnson R.M. and Killeen T. L., pp. 67-87, AGU, Washington, D.C., 1995.

Frederiksen,J. S.: Instability of the three-dimensional distorted polar vortex at the onset of the sudden stratospheric warming, J. Atmos. Sci., 39, 2313-2329, 1982.

Fritts, D. C., and Alexander, M. J.: Gravity wave dynamics and effects in the middle atmosphere. *Rev. Geophys*. 41 (1), 1003, doi:10.1029/2001RG000106, 2003.

Garcia, R. R. and Solomon, S.: A numerical model of zonally averaged dynamical and chemical structure of the middle atmosphere, *J.Geophys. Res.*, **88**, 1379-1400, 1983.

Geller, M. A.: Dynamics of the middle atmosphere, *Space Sci. Rev.* 34, 359-375, 1983.

Gold, E.: The Isothermal Layer of the Atmosphere and Atmospheric Radiation, *Proc. Roy. Soc. Lond.*, 82, 43, 1909.

Hagan, M. E., Forbes, J. M. and Vial, F.: On modelling migrating solar tides, *Geophys. Res. Lett.*, 22, 8, 893-896, 1995.

Hagan, M. E., Chang, J. L. and Avery, S. K.: Global-scale wave model estimates of nonmigrating tidal effects, *J. Geophys. Res.*, 102, 16,439-16,452, 1997.

Hagan, M. E. and Forbes, J. M.: Migrating and nonmigrating semidiurnal tides in the upper atmosphere excited by tropospheric latent heat release, *J. Geophys. Res.*, 108(A2), 1062, doi: 10.1029/2002JA009466, 2003

Harris, M. J., Arnold, N. F. and Aylward, A. D.: A study into the effect of the diurnal tide on the structure of the background mesosphere and thermosphere using the new coupled middle atmosphere and thermosphere (CMAT) general circulation model, *Ann. Geophys.*, 20, 225-235, 2002.

Hartmann, D. L.: Barotropic Instabilities of the Polar Night Jet-Stream, *J. Atmos. Sci.*, vol. 40(4), pp. 817-835, 1983.

Hedin, A. E.: Extension of the MSIS Thermospheric Model into the Middle and Lower Atmosphere, *J. Geophys. Res*. 96, 1159, 1991.

Hernandez, G., Fraser G. J., and Smith R. W.: Mesospheric 12-hour Oscillation near South-Pole Antarctica, *Geophys. Res. Lett.*, 20, 17, 1787-1790, 1993.

Hocking, W. K., Fuller, B., Vandepeer, B.: Real-time determination of meteor-related parameters utilizing modern digital technology, *J. Atmos. Solar-Terr. Phys.*, 63, 155-169, 2001.

Holton, J. R.: An Introduction to Dynamic Meteorology (2nd Edition), Academic Press, 1979.

Holton, J. R.: Wave propagation and transport in the middle atmosphere, *Phil. Trans. R. Soc. Lond.* A 296, 73-85, 1980.

Holton, J. R., *et al*.: *Encyclopaedia of atmospheric sciences*, **2780** pp., Academic Press, Amsterdam; London, 2003.

Jones, M. R.: The dispersion relation for internal acoustic-gravity waves in a baroclinic fluid, *Physics of Fluids*, vol 13 (5), 1274-1280, 2001.

Kato, S.: Non-Migrating Tides, *J. Atmos. Terr. Phys*., 51, 673-682, 2002.

Kingsley, S. and Quegan, S.: Understanding Radar systems, London - McGraw-Hill, 1992.

Kulikov, M. Y.: Theoretical investigation of the influence of a quasi-2-day wave on nonlinear photochemical oscillations in the mesopause region, J. Geophys. Res., 112, D02305, doi:10.1029/2005JD006845, 2007.

Labitzke K., Barnett, J. J. and Edwards B.: Handbook MAP 16, SCOSTEP, University of Illinois, Urbana, 1985.

Lait, L. R. and Stanford, J. L.: Fast, Long-Lived Features in the Polar Stratosphere, *J. Atmos. Sci.*, vol. 45, no. 24, pp. 3800-3809, 1988.

Lieberman, R. S.: Nonmigrating Diurnal Tides in the Equatorial Middle Atmosphere, *J. Atmos Sci.*, 48, 1112-1123, 1991.

Lomb, N. R.: Least-squares frequency analysis of unequally spaced data, *Astrophys. Space Sci.*, 39, 447-462, 1975.

London, J.: Radiative energy sources and sinks in the stratosphere and mesosphere, Proceedings of the NATO Advanced Study Institute on Atmospheric Ozone: Its Variation and Human Influences. Report FAA-EE-80-20, A.Aiken, Ed. 703-721, 1980.

Luo, Y., Manson, A. H., Meek, C. E., Thayaparan, T., MacDougall, J., and Hocking, W. K.: The 16-day wave in the mesosphere and lower thermosphere: simultaneous observations at Saskatoon (52 degrees N, 107 degrees W) and London (43 degrees N, 81 degrees W), Canada, *J. Atmos. Sol.-Terr. Phys*., 64, 1287-1307, 2002.

Malin, S. R. C. and Schlapp, D. M.: Geomagnetic Lunar Analysis by Least-Squares, *Geophys. J. R. Astron. Soc.*, *60*, 409 – 418, 1980.

Manney G. L., Nathan, T. R. and Stanford, J. L.: Barotropic Stability of Realistic Stratospheric Jets, *J. Atmos. Sci.*, vol. 45(18), pp. 2545-2555, 1988.

Manson, A. H., Meek, C. E., Hall, C. M., Nozawa, S., Mitchell, N. J., Pancheva, D., Singer, W. and Hoffmann, P.: Mesopause dynamics from the Scandinavian triangle of radars within the PSMOS-DATAR Project, *Ann. Geophys.*, 22 (2): 367-386, 2004.

McKinley, D. W. R.: Meteor Science and Engineering, McGraw-Hill 1961.

Meek, C. E., Manson, A. H., Franke, S. J., Singer, W., Hoffmann, P., Clark, R. R., Tsuda, T., Nakamura, T., Tsutsumi, M., Hahan, M., Fritts, D. C., Isler, I. and Portnyagin, Yu, I.: Global study of northern hemisphere quasi-2-day wave events in recient summers near 90 km altitude, *J. Atmos. Sol. Terr. Phys.*, 58, 1401-1411, 1996.

Mitchell, N. J., Williams, P. J. S., Beard, A. G., Buesnel, G. R., and Muller, H. G.: Non-linear planetary tidal wave interactions in the lower thermosphere observed by meteor radar, *Annales Geophysicae - Atmospheres Hydrospheres And Space Sciences*, *14*, 364-366, 1996.

Mitchell, N. J., Middleton, H. R., Beard, A. G., Williams, P. J. S., and Muller, H. G.: The 16-day planetary wave in the mesosphere and lower thermosphere, *Ann. Geo. - Atmos. Hydro. And Space Sci.*, *17*, 1447-1456, 1999.

Mitchell, N. J., Pancheva, D., Middleton, H. R., and Hagan, M. E.: Mean winds and tides in the Arctic mesosphere and lower thermosphere, *J. Geophys. Res-Space Phys.*, *107*, 1004(A1), doi:10.1029/2001JA900127, 2002.

Miyahara, S.: The effects on the atmospheric lunar tide of the meridional temperature gradent and the zonal wind, *J. met. Soc. Japan,* 53, 55-68, 1975.

Moudden Y. and Forbes J. M.: Effects of vertically propagating thermal tides on the mean structure and dynamics of Mars' lower thermosphere, *Geophys. Res. Lett.*, 35, L23805, doi: 10,1029/2008GL036086, 2008.

Muller, H. G.: Sheffield Meteor Wind Experiment, *Quart. J. R. Met. Soc.*, 96, 408, 195-213, 1970.

Muller, H. G. and Nelson, L.: A travelling quasi 2-day wave in the meteor region, *J. Atmos. Terr. Phys.*, 40, 761-766, 1978.

Muller, H. G., Havill, R. L., Comley, V. E., and Hill, P. C. J.: A Study Of Meteor Radar Winds From 2 Locations In The British- Isles, *J. Atmos. Terr. Phys.*, *57*, 979-993, 1995.

Murphy, D.J., Tsutsumi, M., Riggin, D. M., Jones, G. O. L., Vincent, R. A., Hagan, M. E. and Avery, S. K.: Observations of a nonmigrating component of the semidiurnal tide over Antarctica, *J. Geophys. Res.*, 108(D8), 4241, doi:10.1029/2002JD003077, 2003.

Nappo, C. J.: An Introduction to Atmospheric Gravity Waves, Academic Press, 2002.

Niu, X. J., Xiong, J. G., Wan, W. X., Ning, B. Q., Liu, L. B., Vincent, R. A. and Reid, I. M., Lunar tidal winds in the mesosphere over Wuhan and Adelaide*, Adv. Space Res.*, 36 (11), 2218-2222, 2005.

Norton, W. A. and Thuburn, J.: The two-day wave in a middle atmosphere GCM, *Geophys. Res. Lett.*, vol. 23, No. 16, 2113-2116, 1996.

Nozawa, S., Imaida, S., Brekke, A., Hall, C. M., Manson, A., Meek, C., Oyama, S., Dobashi, K. and Fujii, R.: The quasi 2-day wave observed in the polar mesosphere, *J. Geophys. Res.*, vol. 108, no. D2, 4039, doi: 10.1029/2002JD002440, 2003a.

Nozawa, S., Iwahashi, H., Brekke, A., Hall, C. M., Meek, C., Manson, A., Oyama, S., Murayama, Y. and Fujii, R.: The quasi 2-day wave observed in the polar mesosphere: Comparison of the characteristics observed at Tromsø and Poker Flat, *J. Geophys. Res.*, vol. 108, no. D24, 4748, doi: 10.1029/2002JD003221, 2003b.

Oberheide, J. and Gusev, O. A.: Observation of migrating and nonmigrating diurnal tides in the equatorial lower thermosphere, *Geophys. Res. Let.*, vol. 29 (24), 2167, doi: 10.1029/2002GL016213, 2002.

Palo, S. E., Roble, R. G., and Hagan, M. E.: Middle atmosphere effects of the quasi-two-day wave determined from a General Circulation Model, *Earth Planets And Space*, *51*, 629-647, 1999.

Palo, S. E., Forbes, J. M., Zhang, X., Russell III, J. M. and Mlynczak, M. G.: An eastward propagating two-day wave: Evidence for nonlinear planetary wave and tidal coupling in the mesosphere and lower thermosphere, *Geophys. Res. Lett.*, vol. 34, L07807, doi: 10.1029/2006GL027728, 2007.

Pancheva, D., Merzlyakov, E., Mitchell, N. J., Portnyagin, Y., Manson, A. H., Jacobi, C., Meek, C. E., Luo, Y., Clark, R. R., Hocking, W. K., MacDougall, J., Muller, H. G., Kurschner, D., Jones, G. O. L., Vincent, R. A., Reid, I. M., Singer, W., Igarashi, K., Fraser, G. I., Fahrutdinova, A. N., Stepanov, A. M., Poole, L. M. G., Maliga, S. B., Kashcheyev, B. L., and Oleynikov, A. N.: Global-scale tidal structure in the mesosphere and lower thermosphere during the PSMOS campaign of June-August 1999 and comparisons with the global-scale wave model, *J. Atmos. Sol.-Terr. Phys.*, *64*, 1011-1035, 2002a.

Pancheva, D., Merzlyakov, E., Mitchell, N. J., Portnyagin, Y., Manson, A. H., Jacobi, C., Meek, C. E., Luo, Y., Clark, R. R., Hocking, W. K., MacDougall, J., Muller, H. G., Kurschner, D., Jones, G. O. L., Vincent, R. A., Reid, I. M., Singer, W., Igarashi, K., Fraser, G. I., Fahrutdinova, A. N., Stepanov, A. M., Poole, L. M. G., Malinga, S. B., Kashcheyev, B. L. and Oleynikov, A. N.: Global-scale tidal variability during the PSMOS campaign of June-August 1999: interaction with planetary waves, *J. Atmos. Sol.-Terr.* Phys., 64 (17), 1865-1896, 2002b.

Pancheva, D., Mitchell, N. J., Manson, A. H., Meek, C. E., Jacobi, C., Portnyagin, Y., Merzlyakov, E., Hocking, W. K., MacDougall, J., Singer, W., Igarashi, K., Clark, R. R., Riggin, D. M., Franke, S. J., Kurschner, D., Fahrutdinova, A. N., Stepanov, A. M., Kashcheyev, B. L., Oleynikov, A. N., Muller, H. G.: Variability of the quasi-2-day wave observed in the MLT region during the PSMOS campaign of June-August 1999, *J. Atmos. Sol.-Terr.* Phys., 66 (6-9), 539-565, 2004.

Plumb, R. A.: Baroclinic Instability of the Summer Mesosphere: A Mechanism for the Quasi-Two-Day Wave?*, J. Atmos. Sci.*, vol. 40, pp. 262-270, 1983.

Portnyagin, Y.I., Forbes, J. M., Makarov, N. A., Merzlyakov, E.G., and Palo, S.: The summertime 12-h wind oscillation with zonal wavenumber s=1 in the lower thermosphere over the South Pole, *Ann. Geo. - Atmos. Hydro. And Space Sci.,16, 7, 828-837,* 1998.

Prata, A. J.: The 4-Day Wave, *J. Atmos. Sci.*, vol. 41, pp. 150-155, 1984.

Randall, D. A.: The Laplace Tidal Equations and Atmospheric Tides, *Selected Notes of David Randall - Department of Atmospheric Science Colorado State University,* 2000.

Riggin, D. M., Meyer, C. K., Fritts, D. C., Jarvis, M. J., Murayama, Y., Singer, W., Vincent, R. A., and Murphy, D. J.: MF radar observations of seasonal variability of semidiurnal motions in the mesosphere at high northern and southern latitudes, *J. Atmos. Sol.-Terr. Phys.*, *65*, 483-493, 2003.

Salby, M. L.: The 2-Day Wave in the Middle Atmosphere: Observations and Theory, *J. Geophys. Res.*, vol. 86, no. C10, pp. 9654-9660, 1981.

Salby, M.L., Hartmann, D.L., Bailey, P.L. and Gille, J.C.: Evidence for equatorial Kelvin modes in Nimbus-7 LIMS, *J. Atmos. Sci*., 41, 220-235, 1984.

Salby, M.L.: Fundamentals of Atmospheric Physics, Academic Press, 61, 1996.

Sandford, D. J., Muller, H. G. and Mitchell, N. J.: Observations of lunar tides in the mesosphere and lower thermosphere at Arctic and middle latitudes, *Atmos. Chem. and Phys*., 6, 4117-4127, 2006.

Sandford, D. J. and Mitchell, N. J.: Lunar tides in the Mesosphere over Ascension Island (8° S, 14.4° W), *Ann. Geophys.,* 25, 9-12, 2007a.

Sandford, D. J., Mitchell, N. J., Vincent, R. A. and Murphy, D. J.: The Lunar tides in the Antarctic mesosphere and lower thermosphere, *J. Atmos. Sol.-Terr. Phys.,* 69, 2219-2237, 2007b.

Sandford, D. J., Schwartz, M. J. and Mitchell, N. J.: The winter two-day wave in the polar stratosphere, mesosphere and lower thermosphere, *Atmos. Chem. and Phys*., 8, 3, 749-755, 2008.

Scargle, J. D.: Studies in astronomical time series analysis: II. Statistical aspects of spectral analysis of unevenly spaced data, *Astrophys. J.*, 263, 835-853, 1982.

Schlapp, D. M. and Harris T. J.: Lunar tidal analysis of 6-yr of mesospheric wind data at Adelaide, *J. Atmos. Terr. Phys.*, 55 (13), 1629-1635, 1993.

Schlapp, D. M., Stening, R. J., Forbes, J. M., Manson, A. H., Meek, C. E., and Vincent, R. A.: N-2 and M(2) lunar tides: Atmospheric resonance revisited, *Ann. Geophys.-Atmos. Hydrospheres Space Sci.*, *14*, 826-836, 1996.

Schoeberl, M. R., Douglass, A. R., Hilsenrath, E., Bhartia, P. K., Beer, R., Waters, J. W., Gunson, M. R., Froidevaux, L., Gille, J. C., Barnett, J. J., Levelt, P. F., and DeCola, P.: Overview of the EOS Aura mission, IEEE Trans. Geosci. Remote Sens., 44, 1066– 1074, 2006.

Schwartz, M. J., Lambert, A., Manney, G. L., Read, W. G., Livesey, N. J., Froidevaux, L., Ao, C. O., Bernath, P. F., Boone, C. D., Cofield, R. E., Daffer, W. H., Drouin, B. J., Fetzer, E. J., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Knosp, B. W., Krüger, K., Li, J.-L. F., Mlynczak, M. G., Pawson, S., Russell J. M. III, Santee, M. L., Snyder, W. V., Stek, P. C., Thurstans, R. P., Tompkins, A. M., Wagner, P. A., Walker, K. A., Waters, J. W., and Wu, D. L.: Validation of the Aura Microwave Limb Sounder temperature and geopotential height measurements, *J. Geophys. Res.*, vol. 113, D15S11, doi:10.1029/2007JD008783, 2008.

Siskind, D. E., Eckermann, S. D. and Summers, M. E.: Introduction - Atmospheric Science across the Stratopause, *Geophysical Monograph* 123, pp1-6, 2000.

Smith, A. K., Pancheva, D. V., and Mitchell, N. J.: Observations and modelling of the 6-hour tide in the upper mesosphere, *J. Geophys. Res.-Atmos.*, *109*, D10105, doi:10.1029/2003JD004421, 2004.

Stening, R. J., Meek, C. E., and Manson, A. H.: Lunar Tidal Winds Measured In The Upper-Atmosphere (78-105 Km) At Saskatoon, Canada, *J. Atmos. Sci.*, *44*, 1143-1151, 1987.

Stening, R. J.: A Diurnal Modulation Of The Lunar Tide In The Upper-Atmosphere, *Geophys. Res. Lett.*, *16*, 307-310, 1989.

Stening, R. J., and Vincent R. A.: A Measurement Of Lunar Tides In The Mesosphere At Adelaide, South-Australia, *J. Geophys. Res-Space Phys.*, *94*, 10121-10129, 1989.

Stening, R. J., Avery, S. K., and Tetenbaum, D.: Observations Of Lunar Tides In Upper-Atmosphere Winds At Poker Flat, Alaska, *J. Atmos. Terr. Phys.*, *52*, 715-721, 1990.

Stening, R. J., Manson, A. H., Meek, C. E., and Vincent, R. A.: Lunar Tidal Winds At Adelaide And Saskatoon At 80 To 100 km Heights - 1985-1990, *J. Geophys. Res-Space Phys.*, *99*, 13273-13280, 1994.

Stening, R., Fleming, K. and Fraser G.: Upper-Atmosphere Semidiurnal tides at Christchurch (44°S) and Scott base (78°S), *J. Atmos. Terr. Phys.*, 57 (8), 857-869, 1995.

Stening, R. J., Forbes, J. M., Hagan, M. E., and Richmond, A. D.: Experiments with a lunar atmospheric tidal model, *J. Geophys. Res.-Atmos.*, *102*, 13465-13471, 1997a.

Stening, R. J., Schlapp, D. M., and Vincent, R. A.: Lunar tides in the mesosphere over Christmas Island (2 degrees N, 203 degrees E), *J. Geophys. Res.-Atmos.*, *102*, 26239-26245, 1997b.

Stening, R. J., Richmond, A. D. and Roble, R. G.: Lunar tides in the thermosphere-ionosphere-electrodynamics general circulation model, *J. Geophys. Res.*, 104, 1-13, 1999.

Stening, R. J., and Jacobi C.: Lunar tidal winds in the upper atmosphere over Collm, *Ann. Geophys.-Atmos. Hydrospheres Space Sci.*, *18*, 1645-1650, 2001.

Stening, R. J., Tsuda, T., and Nakamura, T.: Lunar tidal winds in the upper atmosphere over Jakarta, *J. Geophys. Res-Space Phys.*, *108(A5)*, 1192, doi:10.1029/2002JA009528, 2003.

Swinbank, R. and O'Neill, A.: A stratosphere-troposphere data assimilation system, *Monthly Weather Review*, Vol. 122, 686-702, 1994.

Teitelbaum, H., and Vial, F.: On Tidal Variability Induced by Non-linear Interaction with Planetary Waves, *J. Geophys. Res.,* 96(A8), 14169-14178, doi:10.1029/91JA01019, 1991.

Tsuda, T., Tanii, J., Aso, T., and Kato, S.: Lunar Tides At Meteor Heights, *Geophys. Res. Lett.*, *8*, 191-194, 1981.

Vegard, L.: Results of Investigations of the auroral spectrum during the years 1921-1936, Geophys. Publ, 9(11), 1-71, 1932.

Venne, D. E. and Stanford, J. L.: An Observational Study of High-Latitude Stratospheric Planetary-waves in Winter, *J. Atmos. Sci.*, vol. 39(5), pp. 1026-1034, 1982.

Verne, J.: De la Terre à la Lune, London - Granada Publishing, 1865.

Vial, F., and Forbes J. M.: Monthly Simulations Of The Lunar Semi-Diurnal Tide, *J. Atmos. Terr. Phys.*, *56*, 1591-1607, 1994.

Vincent, R. A.: MF/HF radar measurements of the dynamics of the mesopause region – A review, *J. Atmos. Sol. Ter. Phys*., Vol. 46, No. 11, pp. 961-974, 1984.

von Zahn, U., Hoffner, J., Eska, V., Alpers M.: The mesopause altitude: Only two distinctive levels worldwide?, *Geophys. Res. Lett*, 23, 3231, 1996.

von Zahn, U.: Are Noctilucent Clouds Truly a “Miner’s Canary” for Global Change?, *EOS, Transactions, American Geophysical Union*, 2003.

Wallace, J. M. and Hobbs, P. V.: Atmospheric science: an introductory survey, second edition, London - Academic Press, 2006.

Waters, J. W., Froidevaux L., Harwood R. S., Jarnot R. F., Pickett H. M., Read W. G., Siegel P. H., Cofield R. E., Filipiak M. J., Flower D. A., Holden J. R., Lau G. K., Livesey N. J., Manney G. L., Pumphrey H. C., Santee M. L., Wu D. L., Cuddy D. T., Lay R. R., Loo M. S., Perun V. S., Schwartz M. J., Stek P. C., Thurstans R. P., Boyles M. A., Chandra K. M., Chavez M. C., Chen G., Chudasama B. V., Dodge R., Fuller R. A., Girard M. A., Jiang J. H., Jiang Y., Knosp B. W., LaBelle R. C., Lam J. C., Lee K. A., Miller D., Oswald J. E., Patel N. C., Pukala D. M., Quintero O., Scaff D. M., Van Snyder W., Tope M. C., Wagner P. A. and Walch M. J.: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura Satellite, *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1075-1092, 2006.

Winch, D. E., and Cunningham R. A.: Lunar Magnetic Tides at Watheroo: Seasonal, Elliptic, Evectional, Variational and Nodal, *Journal of Geomagnetism and Geoelectricity*, *24*, 381 – 414, 1972.

Winch, D. E.: Spherical harmonic-analysis of geomagnetic tides, 1964-1965, *Phil Trans. R. Soc. Lond.*, 303, 1-104, 1981.

Wu, D. L., Hays, P. B. and Skinner, W. R.: A Least Squares Method for Spectral Analysis of Space-Time Series, *J. Atmos. Sci.*, vol. 52, no. 20, pp. 3501-3511, 1995a.

Wu, D. L., Fishbein, W. G. and Waters, J. W.: Excitation and Evolution of the Quasi-2-Day Wave Observed in UARS/MLS Temperature Measurements, *J. Atmos. Sci.*, vol. 53, no. 5, pp. 728-738, 1995b.

Younger, P. T., Pancheva, D., Middleton, H. R., and Mitchell, N. J.: The 8-hour tide in the Arctic mesosphere and lower thermosphere, *J. Geophys. Res-Space Phys.*, *107(12A)*, 1420, doi:10.1029/2001JA005086, 2002.